TECHNICAL UNIVERSITY OF CRETE
School of Electrical and Computer Engineering

Developing an Autonomous Agent for Automated Electricity Trading

Diploma thesis

Stavros Orfanoudakis

Thesis Committee:
Advisor: Georgios Chalkiadakis, Associate Professor ECE TUC
Fotios Kanellos, Associate Professor ECE TUC
Vasilis Samoladas, Associate Professor ECE TUC

Chania, February 2021

Acknowledgements

I would like to especially thank my supervisor Prof. Georgios Chalkiadakis as well as Charilaos Akasiadis, who was the other mentor of the TUC-TAC team, for their consistent support and guidance during the running of this project.

Abstract

The rise of renewable energy production, along with the recent popularization of electric vehicles, are gradually creating the need for a smarter electricity grid. Against this background, the PowerTAC trading agents competition and platform offers researchers (from academia and the industry alike) with effective means to test different business, market analysis, and market prediction strategies (potentially along with novel artificial intelligence algorithms), before even deploying them in the Smart Grid. In more detail, PowerTAC constitutes a multi-agent simulation platform for electricity markets, in which intelligent agents corresponding to electricity brokers compete with each other aiming to maximize their profits. Now, as AI researchers have found out the hard way time and time again, greediness almost never pays off in competitive multi-agent settings. In PowerTAC, too, agents that aim to take over a disproportionately high share of the market, might end up incurring financial losses due to being obliged to pay huge transmission capacity fees. Starting from this observation, we developed a novel trading strategy that aims to balance gains from controlling a sufficiently large part of the retail market, against the costs of paying high transmission capacity fees. We equipped TUC-TAC 2020, an agent that represented the Technical University of Crete in the PowerTAC-2020 international competition with this retail market strategy. Moreover, we developed a wholesale market strategy that utilized Monte Carlo Tree Search to determine TUC-TAC's best course of action when participating in the market's double auctions. Using these strategies, TUC-TAC was crowned the PowerTAC-2020 champion, competing against 7 other agents representing universities from 6 different countries. In this thesis, we present TUC-TAC's 2020 strategy in detail; and also conduct an extensive post-tournament analysis, in order to draw important lessons regarding the strengths and weaknesses of the various strategies used in the PowerTAC-2020 competition.

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Ανάπτυξη Αυτόνομου Πράκτορα για Αυτοματοποιημένες Αγοραπωλησίες Ηλεκτρικής Ενέργειας

Διπλωματική Εργασία

Σταύρος Ορφανουδάκης

Εξεταστική Επιτροπή:

Επιβλέπων: Γεώργιος Χαλκιαδάκης, Αναπληρωτής Καθηγητής Η.Μ.Μ.Υ. Π.Κ.

Φώτιος Κανέλλος, Αναπληρωτής Καθηγητής Η.Μ.Μ.Υ. Π.Κ.

Βασίλης Σαμολαδάς, Αναπληρωτής Καθηγητής Η.Μ.Μ.Υ. Π.Κ.

Χανιά, Δεκέμβριος 2020

Περίληψη

Η μεγάλη αύξηση της παραγόμενης από ανανεώσιμες πηγές ενέργειας, και η συνεπακόλουθη ένταξή της σε μεγάλη κλίμακα στην οικιακή αγορά ενέργειας, παράλληλα με την εξελισσόμενη διάδοση χρήσης των ηλεκτρικών οχημάτων, δημιουργούν σταδιακά την ανάγκη για ένα πιο «έξυπνο» ηλεκτρικό δίκτυο - το λεγόμενο Smart Grid, ή Έξυπνο Δίκτυο Ηλεκτροδότησης (ΕΔΗ) στα ελληνικά. Σε αυτό το διαμορφούμενο περιβάλλον, ο διαγωνισμός PowerTAC (Power Trading Agent Competition) προσφέρει τη δυνατότητα σε ερευνητές (προερχόμενους τόσο από πανεπιστήμια όσο και από τη βιομηχανία) να δοκιμάσουν διάφορες στρατηγικές - επιχειρηματικές, ανάλυσης ή πρόβλεψης της αγοράς (πιθανότατα και νέων αλγορίθμων Τεχνητής Νοημοσύνης) - προτού τις εφαρμόσουν στο ΕΔΗ. Πιο συγκεκριμένα, το PowerTAC αποτελεί κατά κύριο λόγο μία πλατφόρμα ανάπτυξης και δοκιμής έξυπνων πρακτόρων λογισμικού που ανταγωνίζονται ως πάροχοι ενέργειας στις διάφορες αγορές του (μελλοντικού) ΕΔΗ, στοχεύοντας στην μεγιστοποίηση των κερδών τους. Βέβαια, πλείστες όσες ερευνητικές εργασίες και εφαρμογές Τεχνητής Νοημοσύνης έχουν επανειλημμένα καταδείξει το γεγονός ότι οι άπληστες στρατηγικές σχεδόν ποτέ δεν αποδίδουν σε ανταγωνιστικά περιβάλλοντα πολλαπλών πρακτόρων. Ομοίως στο PowerTAC, οι πράκτορες που στοχεύουν να αποκτήσουν ένα δυσανάλογα υψηλό μερίδιο της αγοράς, ενδέχεται να καταλήξουν σε οικονομικές απώλειες λόγω της υποχρέωσης καταβολής τεράστιων σε μέγεθος τελών στις ρυθμιστικές αρχές. Ξεκινώντας από αυτήν την παρατήρηση, σε αυτή την εργασία αναπτύξαμε μια καινοτόμα στρατηγική εμπορικής διαπραγμάτευσης, που στοχεύει στην εξισορρόπηση των κερδών ελέγχοντας ένα αρκετά μεγάλο μέρος της αγοράς, έτσι ώστε να αντισταθμίσει το υψηλό κόστος των τελών. Εξοπλίσαμε τον TUC-TAC 2020, τον ευφυή πράκτορα λογισμικού που εκπροσώπησε το Πολυτεχνείο Κρήτης στο διεθνή διαγωνισμό PowerTAC-2020, με αυτήν τη στρατηγική για χρήση στην λιανική αγορά ενέργειας του διαγωνισμού. Επιπλέον, ο πράκτορας μας χρησιμοποιεί Δενδρική Αναζήτηση Monte Carlo για να εντοπίσει το καλύτερο σχέδιο δράσης και να καθορίσει τις προσφορές του κατά την συμμετοχή του σε χρηματιστηριακού τύπου δημοπρασίες (double auctions) στη χονδρική αγορά ενέργειας του διαγωνισμού. Χρησιμοποιώντας αυτές τις στρατηγικές, ο πράκτορας TUC-TAC στέφθηκε πρωταθλητής του PowerTAC-2020, αντιμετωπίζοντας 7 αντιπάλους που εκπροσωπούσαν πανεπιστήμια από 6 διαφορετικές χώρες. Σε αυτήν την διπλωματική εργασία, παρουσιάζεται λεπτομερώς η στρατηγική του TUC-TAC 2020, και επιπλέον παρέχεται μια εκτενής ανάλυση των αποτελεσμάτων του διαγωνισμού, με σκοπό την άντληση σημαντικών μαθημάτων σχετικά με τα πλεονεκτήματα και τις αδυναμίες των διαφόρων στρατηγικών που χρησιμοποιήθηκαν στον διαγωνισμό PowerTAC-2020.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter we present the motivation for our work, we also outline our approach and main contributions, as well as, provide an overview of the rest of this thesis.

1.1 Motivation

The rise of renewable energy production in the residential market along with the latest popularization of electric vehicles is gradually creating needs for a "smarter" grid. The necessity of this new Grid is indisputable because of the unique features it will be consisted of. In Smart Grid settings, one of the main purposes is to reduce fossil fuel consumption. This is especially important for two main reasons. The first reason is that fossil fuel sources will be depleted at some point in the future, so alternative energy sources will be eventually required. The other reason is the bad impacts to the planet's physical environment that subsequent coal emissions from the burning of fossil fuels have.

So, a feature of the Smart Grid will be an energy market which will consist of a lot of more participants, with most of them being able to buy and sell energy at the same time, and indeed doing so. Hence, researchers need a tool that will help them to experiment in novel ways to make this new market viable. The Power Trading Agent Competition(PowerTAC) is a rich simulation platform that can provide researchers with efficient ways to try and test different strategies and approaches before actually deploying them in the future Smart Grid. The PowerTAC platform right now has the most features a smart electricity grid can possess (interruptible consumption, electric vehicles, renewable energy, and so on) so the simulations can be as realistic as possible.

The main objective of our research was to create an autonomous intelligent agent that will be able to adapt to the complex and highly competitive energy markets of the PowerTAC with the purpose of making a profit. In order to implement our agent, TUC-TAC, we tested many state of the art implementations of other already competing brokers and in the end, we chose the strategies that were more suitable for our purposes. The overall TUC-TAC implementation was split into two parts. The first part which is extensively discussed in this thesis regards all the processes of the decision-making of our agent in the retail and wholesale market. The other part was about creating the necessary predictors which would help in some of the aspects of our agent and is considered as a separate module of the overall agent. So, our team consists of Stavros Orfanoudakis, Stefanos Kontos, Georgios Chalkiadakis, and Charilaos Akasiades.

The base principle that we applied has certain analogies to the equilibrium strategy used by the winning agent of the 2010 Lemonade Stand Game tournament [2]. In a few words, their winning strategy was to try to coordinate with an opponent and sit opposite of him so their agent could have the highest utility at all times. Our strategy is quite similar to that since its basic goal is to get half the available market share leaving the rest to the others. By doing that, the TUC-TAC agent expects to always have the highest income, while sharing the fees with the other agents. We also implemented the Monte Carlo Tree Search for bidding in the double auction of the wholesale market, similar to that developed by Chowdhury et al [3].

1.2 Overview of the Thesis

The work described in this thesis represents the core of the TUC-TAC's strategy and in particular the agent's strategy for participating in the retail and wholesale market of PowerTAC. The rest of this thesis is structured as follows: In Chapter 2 we present the state of the art technologies and ideas that can make the existence of the Smart Grid possible. Following the theoretical background of Chapter 2, in Chapter 3 we describe in detail the rules and the components of the PowerTAC competition and present the current PowerTAC state of the art implementations of our opponents. Afterward, in Chapter 4 we describe in detail every aspect of the TUC-TAC 2020 agent. Finally, in Chapter 5, we outline the early development and the preparation of the TUC-TAC agent, so then we finally discuss about our agent's performance in the PowerTAC 2020 competition, and also observe the results of extensive post-tournament analysis.

Chapter 2

Background

As time passes and the natural reserves of fossil fuels are getting depleting, the need for an alternative way to produce energy emerges. Even more so, since fossil fuels contribute greatly to CO2 emission leading to climate change with all its consequences. Currently, there are many ways to generate electrical energy without burning fossil fuels by using renewable energy resources. These are harvested via solar panels, wind turbines, tidal generators and other emerging technologies. As many researchers (Amin et al [4]) have already pointed out, in order to create an energy grid that will be solely supported by renewable energy sources, many changes must be made. Thus, the creation of a new, Smart Grid is necessary to put all of these practices into effect. Some of the features and concepts of the Smart Grid [5] that will be discussed in this Chapter are: Energy Prosumers, Virtual Power Plants, Demand Side Management and at last Electric Vehicles and the G2V/ V2G problem.

2.1 Energy Prosumers and Virtual Power Plants

One key aspect of the emerging Smart Grid is the ability of each participant to produce its own energy and contribute to the network. Thus, potentially a lot of individual homes and businesses will eventual become "energy prosumers". In some detail, an energy prosumer is a consumer who also produces renewable energy, for example by using photo-voltaic panels at the rooftop of his home or his shop. By doing that, the stock prices of the electricity will drop significantly while the available energy will be increased. In addition to that, prosumers can form teams which in a large scale can also help in shifting the energy load of the grid [6], thus reducing, even more, energy waste. One example of artificial intelligence technologies that can help in solving the "prosumers" problem is described in the work of Angelidakis et al [7], [8]. Specifically, they propose factored MDPs for Optimal prosumer decision-making in continuous state spaces, which can help in modeling the complex state space of the Smart Grid, where prosumers produce and consume energy simultaneously.

Eventually, at some point in the Smart Grid, many individual "prosumers" or simple consumers will exist with each one of them trying to maximize its profit. So, in order to comply with the demand-side management aspect of the grid, it will be necessary for these individuals to create coalitions, so they can balance the supply and demand in a part of the network, acting as a Virtual Power Plant (VPP). Specifically:

> "The process of forming VPPs at a technical level means the individual actors must synchronize the largely heterogeneous services they provide within the VPP in an agile fashion to meet the requirements of the contracts they make with their customers."

: VPP definition taken from the survey of Ramchurn et al. [9]

Moreover, Chalkiadakis et al [10] and Robu et al [11] discuss the issues surrounding the emergence of Virtual Power Plants, and propose game-theoretic and mechanism design solutions to help tackle the inefficiency and unreliability problems plaguing the integration of distributed energy resources into the Smart Grid. That research highlights the need to create decentralized coordination algorithms and strategies that will enable the creation of efficient coalitions of distributed energy resources (i.e., VPPs) in finite time.

2.2 Demand-Side management

Unlike the current energy grid state where electricity flows only from big power facilities to the consumers, in the new Smart Grid electricity will flow in any direction in different magnitudes. This statement is in the core of the demand side balancing problem, whose solution is necessary in order to reduce any energy waste. Effective demand-side balancing will benefit consumers who shift their consumption to desirable time slots, in order to enjoy cheaper tariffs. Currently, the main approach to reduce consumers' demand is to directly control the high load electrical devices by shutting them down when the energy demands of the grid cannot be fulfilled. By contrast, in the Smart Grid every household, shop, or industrial site will communicate, and receive signals from the grid with the purpose of managing the energy demand.

However, in order to effectively deploy this idea, more complicated tariffs must be offered to cover the wide variety of customers along with their unique needs. To do so, it is necessary to create mathematical model simulation systems that will precisely show us how the consumers and the grid react in different scenarios; and which will allow the application of novel mechanisms, which will provide incentives to prosumers to produce and/ or consume energy when required. This will be key in order to keep the Grid balanced and to ensure the optimal use of renewable energy sources. So in order for this problem to be properly researched, it can be turned into an online learning and management problem under uncertainty, the solution of which can benefit from game theoretic ideas. For example, Akasiadis and Chalkiadakis [6] demonstrate that demand-side management can be greatly benefited when consumer agents form cooperatives to effectively shift the power consumption. Also, Kota et al [11] propose mechanisms to enable the formation of consumer cooperatives and facilitate demand-side management: the idea is that a cooperative would recruit suitable electricity consumers as members, who agree to participate by attempting to reduce their energy consumption when requested.

2.3 Electric Vehicles and the G2V/ V2G problem

As it was previously mentioned, one of the goals of the Smart Grid is to reduce coal emissions, so standard vehicles should be replaced by electric vehicles. At this time, car manufacturers are already adapting to the new needs and develop EVs aiming to sell them to the average consumer. So it is highly possible that in the near future the majority of the cars will be powered solely by electricity. Specifically, in the work of Jenn et al [12] is mentioned that in 2015 the market share of EVs in the more advanced states of U.S.A. was already up to 3% and it is projected to increase exponentially in the coming years.

Aside from the reduction of the harmful emissions that EVs can accomplish, EVs can also help in demand-side management because of the huge batteries they carry. Specifically, EVs will be charged when the energy is plentiful and will be discharging back to the Grid when it needs a portion of the energy back (Vehicle-to-Grid) [13] to ensure its stability. Meanwhile, EVs store electricity which is used to power the car too , thus it is necessary for these batteries to be recharged at some point (Grid-To-Vehicle). However from the Grid's perspective, charging EVs greatly impact the state of the network because an increased number of simultaneous battery charges adds a significant load to it [5]. Thus, it is imperative to come up with efficient, coordinated charging schemes that guarantee the smooth Grid operation [13] [14]. On the other hand the greatest challenge for V2G is the uncertainty of the power that can be supplied by EVs because of their location and needs. For all these reasons, it is obvious that an increased EVs penetration level into the Grid, makes the need to take them into careful consideration when designing consumption and production prediction models imperative.

Chapter 3

PowerTAC: The Power Trading Agent Competition

After our discussion in the previous Chapter about the Smart Grid and its components, the necessity for a platform to exist that would enable researchers to simulate in real-time some of the features of the future smart grid is apparent. Hence, Ketter and Collins[15] with their team created the PowerTAC platform. PowerTAC is a rich competitive economic simulation of future energy markets, featuring several smart grid components. With the help of this simulator, researchers are able to better understand the behavior of future customer models as well as experiment with retail and wholesale market decision-making, by creating competitive agents and benchmark their unique strategies against each other. In this way, a host of useful information is extracted which can be used by policymakers and industries in order to prepare for the upcoming market changes.

3.1 Competition Overview

PowerTAC was first developed in 2011, and since then every year the developers manage to update some of the components like customer models in order to increase the accuracy of the simulation as well as to catch up with the latest scientific research. The PowerTAC vision consists of competitive agents that will harness the energy supply and demand of the simulation environment with the intention of making a profit. Specifically the "broker"-agents will buy and sell energy through consumption tariffs with individual retail customers like households, commercial stores, or even bigger enterprises as well as electric vehicles. At the same time, the agents will interact and trade within the wholesale market which is a real-world replica of the European and North American wholesale energy markets [16].

As the competition organizers mention, this simulation is designed to model the energy trading environment mainly from an economic and not an especially technical viewpoint. Through the years, the *main* elements of PowerTAC have not changed and can be seen in Figure 3.1 below. More details about these components will be further discussed in the following sections.

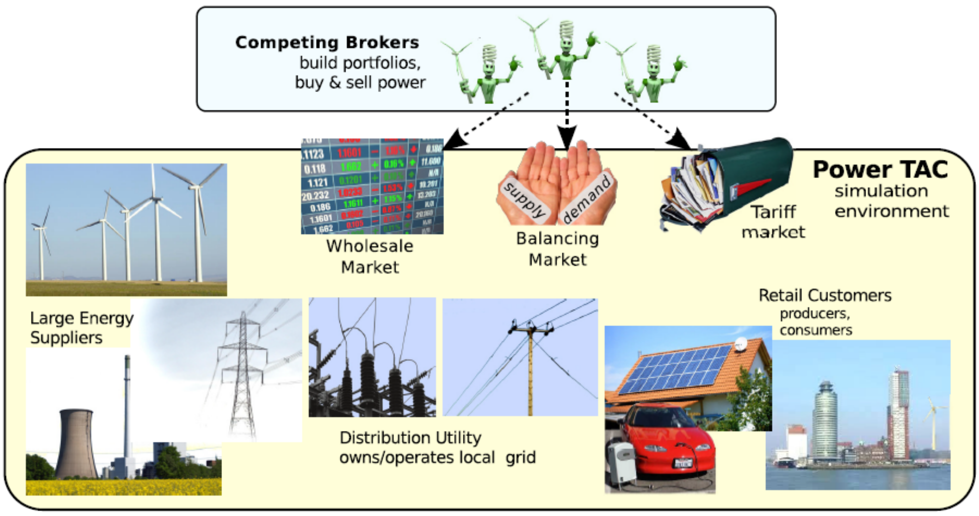

Figure 3.1: Main PowerTAC elements [1]

3.1.1 Simulation Time

For a simulation to work, the game time needs to be discrete, thus some discrete time-blocks are created. These are called "time slots" and each one of them represents an hour of simulated time while each one of them takes almost 5 seconds of real-time. Each game consists of at least 1440 timeslots (two months of simulated time and at least 2 hours of real-time per game). This means that in each PowerTAC game at any time there is an active timeslot, and a set of future timeslots for which the brokers can reserve and trade energy. The main objective of the brokers is to try to balance the demand and supply in each of the future timeslots with the purpose of not getting monetary penalties.

3.1.2 Brokers

Broker-agents are the real-life analogy to energy retailers and have similar actions in their activity book. In each timeslot, every agent can decide and perform any of the actions seen in Figure 3.2 below.

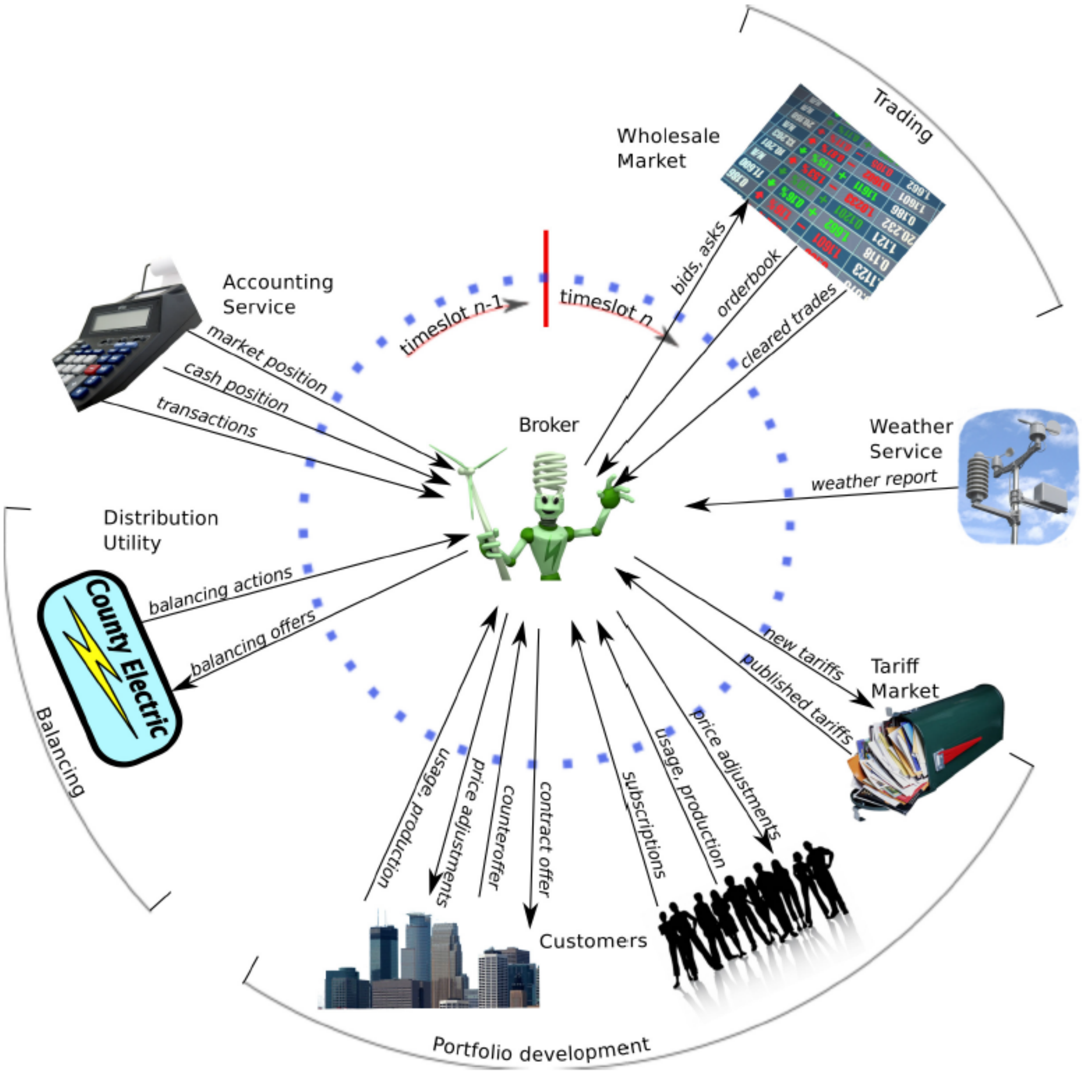

Figure 3.2: Available actions of a broker during a timeslot[1]

Specifically these actions are:

- **Publish new tariffs in the retail market.**
- **Modify existing tariff terms** by revoking the old tariff and publishing a new one in its place.
- **Adjust tariff prices**, if tariff terms allow it.
- **Trade in Wholesale Market** by placing bids or asks to sell or procure energy for future timeslots.
- **Curtail demand** by issuing economic control orders which apply to customers with incorruptible consumption tariffs.
- **Submit balancing orders in balancing market** which consists of offering controllable capacities for actual time balancing.

However, there is a lot of information available to the brokers in order to help them make the correct decisions. In a few words agents know about:

- **Most of the game parameters**, such as multipliers, cost specific parameters which are stable through the course of a game and are essential to calculate some values in advance.
- **Broker identities**, only the names of the competitor agents.
- **Bootstrap data**, consists of the net demand and the electricity usage of each customer as well as the bid and ask prices that are cleared in the Double Auction of the Wholesale Market through the duration of the bootstrap period. The bootstrap period is a time in which only the default broker is active and is used to initialize a game. The Bootstrap data is usually used by the agents to generate accurate electricity usage customer models.

- **Weather reports/forecasts**.
- **Active Tariffs of each broker**.
- **Wholesale market clearing data**.
- **Wholesale market order-books**.
- **Total aggregate energy consumption/production**.
- **Every Transaction is private data for each agent**, like cash position, market positions, portfolio supply demand, balancing and distribution transactions.

3.1.3 Tariffs in Retail Market

Tariffs are the main and necessary way for an agent to make a profit. In PowerTAC there are specific tariffs for a wide range of different power consumers/producers types. Some of these are general consumption, interruptible consumption, general storage, battery storage, thermal storage consumption, electric vehicles, general production, wind production, solar production, and some other similar power type not currently used by the game server in practice.

The PowerTAC tariff creation gives the option to brokers to create and publish tariffs with each possible term deriving from the needs of current but also future energy markets. In Figure 3.3 below we can see the current tariff structure.

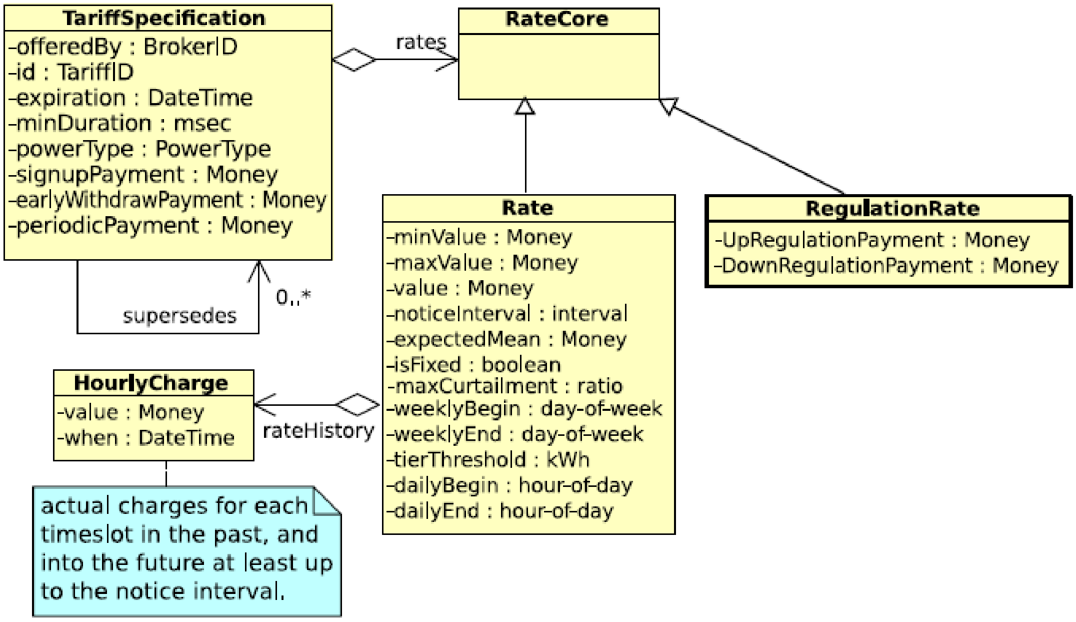

Figure 3.3: Detailed Tariff Parameters [1]

It is easily visible that with these tariff terms countless different tariffs can be offered. Some of the common tariff features that are supported are:

- **Tiered Rates**, which mean that customers can be charged with more than one rate for electricity: a lower rate for the electricity used up to a certain threshold; and a second, higher rate for all additional use.
- **Time of Use rates**, which is a method of measuring and charging a customer's energy consumption based on when the energy is used.
- **Two-part tariffs** are a combination of fixed daily fee plus usage fee.
- **Sign up and Early withdrawal penalties**, these features can be used as a clause of a tariff.
- **Interruptible Rates**, which means that the tariff specifies the demand level (kW) that will be interrupted. To use this option, the customer must

have on-site generation or storage, to serve the load they have nominated for interruption.

- **Regulation Rates** are used by storage devices and can specify different prices for the use of the device for up-regulation or down-regulation.

- **Dynamic Pricing tariffs** in which the new prices must be communicated to subscribed customers some number of timeslots before the timeslot to which they apply.

3.1.4 Customer Models

The most important and the most updated feature of PowerTAC is that of the customer models. Customer models are the ones responsible for interacting with the retail market in order to find the best tariffs as well as generate the values for the energy consumption or generation during a timeslot taking into account many different variables like weather, market state, etc. Each customer model is defined by its name, population, power type, controllable capacity, and multi-contracting. The customers' models themselves are of high complexity so these could be the main focus of other scientific researches, therefore no more technical details will be presented in this work. More details can be seen in the corresponding chapter of PowerTAC definition [1].

3.1.5 Weather Reports

Another significant feature of PowerTAC is the weather reports. In each timeslot, a weather report for the current time is sent to the brokers along with a weather forecast about the next 24 timeslots. This information is used by the agents for prediction purposes, because the customer models are directly affected by the weather. Specifically, weather data are drawn from real weather reports in the past which means that, during the competition agents do not know the game location for obvious reasons.

3.1.6 Wholesale Market

The wholesale market is the place that allows brokers to primarily buy and sell energy for every available future timeslot (usually 24 future timeslots are enabled). Specifically, the PowerTAC wholesale market is a "day ahead" periodic double auction, which clears in each timeslot and tries to imitate the current core wholesale trading foundations like FERC, EEX, and NordPool (wholesale trading utilities in Europe and North America).

In real life, energy wholesale markets serve big regions consisting of many energy providers and millions of customers. On the other hand, the PowerTAC can currently simulate only one city from such an area. Therefore, in order to make the *wholesale market* more realistic, three more entities were created to trade in that market along with the brokers.

The first entity is called "Grid Genco" and represents the wide population of generating facilities that can supply the simulated city. This entity is very important because it provides the market with a realistic supply curve deriving from statistics observed in MISO and PJM LMP markets of North America. The second entity is the "Grid Buyer" and simulates the regional demand also based on real-life metrics drawn from a time series trained on two years of MISO North actual demand. So, this entity's responsibilities are to buy enough energy so it can satisfy its energy demands. The third and last entity is created to provide the market with liquidity, using a stochastic behavior that mimics a population of buyers and speculators, which are only engaging in the Wholesale Market of PowerTAC and not in the Retail Market.

3.1.7 Balancing Market and Balancing Fees

The balancing market is the real-life equivalent of Independent System Operator (ISO) in the U.S. and Transmission Systems Operator (TSO) in Europe. Its main responsibilities are to monitor the electricity grid and maintain balance by keeping voltage and frequency within some bounds. However, in PowerTAC technical aspects like voltage and power factor are not taken into consideration so the balancing market's only responsibility is to balance the supply and demand in each timeslot by exercising capacity controls on behalf of the agents. In this way, when a broker fails to procure the required energy in time, balancing utility comes and finally charges the broker for the missed energy at a much higher price, acting as a penalty. That fee is called *Balancing Fee* and can vary from smaller charges to very high destructive penalties depending on the Wholesale market prices.

3.1.8 Distribution utility and Transmission Capacity Fees

The Distribution utility (DU) is primarily responsible for 3 different operations. The first one, as its name suggests, is to distribute the energy to each customer while it charges each broker with the distribution fees which are relevant to the energy transmitted through the grid. Second, DU is responsible to issue the Transmission Capacity Fees (TSF). These fees represent the amount of money a broker should pay for its customers' contribution to demand peaks. This means that when there is a demand peak each broker will have to pay for a portion of the exceeding energy. TSF charge the three highest demand peaks at the end of a 168 timeslots period (1 week of simulated time). Specifically the calculation of the TSF consists of three steps:

- Compute the mean d_t and standard deviation $\sigma_{d,t}$ of the net demand since the start of the game, and define threshold as $z_t = d_t + \gamma\sigma_{d,t}$ where γ is a tournament constant.

- Find the three highest net demand timeslots.
- For each of these three timeslots if demand p is higher than the threshold the total charge will be $\Phi = \lambda(p - z_t)$ and λ is a tournament constant.

In the current PowerTAC competition these fees are the main problem an agent faces when it tries to dominate in the retail market. More details about how our agent, TUC-TAC, managed to harness these fees will be presented in the next chapter that describes our approach. The last responsibility of the Distribution utility is to publish some default tariffs for the times when there are no published tariffs by any other broker.

3.2 Related Work

PowerTAC as mentioned earlier is almost 10 years old. Specifically, in 2012 the first pilot tournament had taken place with 8 research teams from around the world and since then it is organized every year. Having so many years of active competition implies, that there are many different interesting agent approaches already implemented. In this section, some of the most significant broker-agent strategies will be introduced. Every agent design, in these many years of competition, can be broadly separated into two different, almost autonomous, parts. The first part is the Retail Market Module which in general tries to find the best tariff strategy and the second is the wholesale market module which its main responsibilities are to submit bid and asks in the periodic double auction.

3.2.1 Retail Market Strategies

Many agents in the past, like ColdPower2017 agent used Markov Decision Processes (MDP) [17] in order to find the best tariff strategy. Also, other agents tried similar strategies, one of them is Mertacor2020 which implements Q-Learning techniques in order to maximize the profits from the retail market. VidyutVanika [18] Agent also used a combination of MDP and Q-learning assisted by a Deep Neural Network Predictor. However, AgentUDE [19], one of the most successful agents in the many years of PowerTAC, which won the tournaments of 2014, 2017, 2018 and was in the top three brokers in 2016, and 2019, used a much simpler tariff strategy. Specifically, its strategy was based mainly on decision trees and it was being enhanced with some general principles; for example, the name of such a principle was "revokeUselessTariffs", which meant to revoke all potentially harmful tariffs, and so on. In addition to that strategy, AgentUDE2017 [20] added a genetic algorithm module to further improve its tariff generation.

3.2.2 Wholesale Market Strategies

The double auction of the wholesale market requires a very careful strategy in order to be profitable. One of the first and most important works in this field was that of TacTex [21] agent in 2013. That team used an MDP price predictor which is the foundation of almost all modern brokers in PowerTAC. Specifically, SPOT [3] agent further improved the previous strategy using Monte Carlo Tree Search to find the best bids and asks at the best possible times. Another especially efficient wholesale market agent was VidyutVanika [18], which also uses the MDP based price predictor which was firstly implemented by TacTex 2013. In all these years there were also many other interesting works like that of Nativdad et al(2016) [22] which was using machine learning techniques to reduce the complexity of the wholesale market action space.

Chapter 4

Our Approach

The TUC-TAC agent is a complex artificial intelligence software developed to compete in the 2020 Power Trading Agents Competition (PowerTAC-2020). The main strategy of TUC-TAC 2020 (more specifically, the part of TUC-TAC's strategy that is used in the key for the game retail market), is based on the principle that, acquiring half of the market share will give TUC-TAC half of the total profits, but also only half of the inevitable transmission capacity fees will have to be paid by our agent. We came up with that resolution after realizing that greedy strategies would not work in the competitive environment of PowerTAC; and we were inspired in the choice of strategy by an interesting equilibrium strategy employed in the context of the "Lemonade Game" competition, and which we briefly present in the next section. In order to achieve that, the TUC-TAC uses decision trees enhanced with many heuristics and non-heuristics functions that help in the evaluation of the game state. We also employed the Monte Carlo Tree Search algorithm for bidding in the double auction of the wholesale market, adapting it to this setting. In this chapter, we will break down the agent into modules to easier understand how it was designed. Figure 4.1 below depicts the main components of the agent; in this chapter, we will analyze all these components in turn.

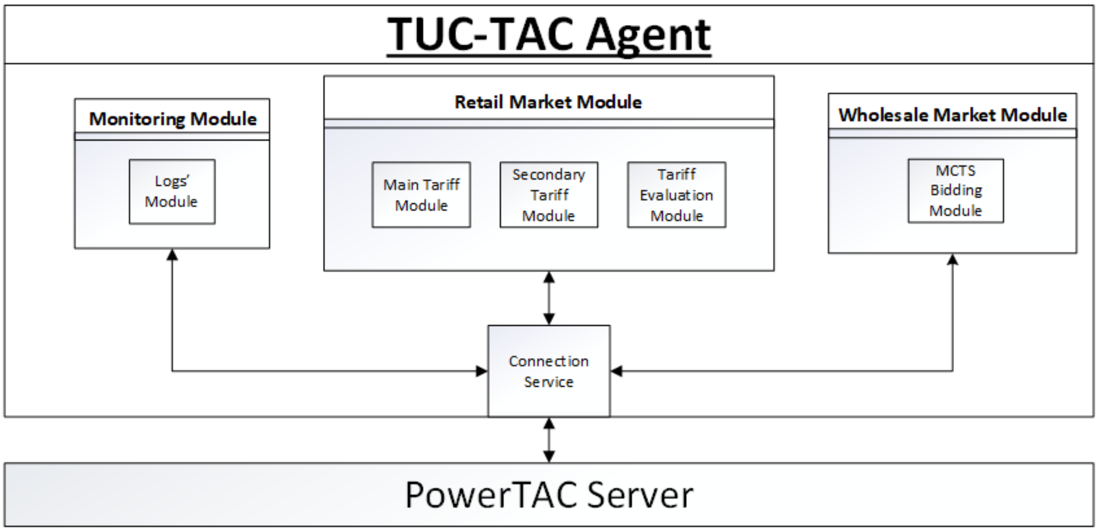

Figure 4.1: TUC-TAC's architecture

4.1 Monte Carlo Tree Search

Before we turn into describing the various components of our agent's strategy, we begin this chapter by providing some background on the Monte Carlo Tree Search (MCTS) technique, which was employed in our wholesale market strategy, as we explain later in this chapter. In general, Monte Carlo Tree Search (MCTS) is a search technique in the field of Artificial Intelligence. It is a probabilistic and heuristic driven search algorithm that combines the classic tree search implementations alongside machine learning principles of reinforcement learning.

The basic algorithm involves iteratively building a search tree until some predefined computational budget —typically a time, memory, or iteration constraint— is reached, at which point the search is halted and the best performing root action is returned. Each node in the search tree represents a state of the domain, and directed links to child nodes represent actions leading to subsequent states [23].

- 1. *Selection*: Starting at the root node, a child selection policy is recursively applied to descend through the tree until the most urgent expandable node is reached. A node is expandable if it represents a nonterminal state and has unexpanded.

- 2.*Expansion*: One (or more) child nodes are added to expand the tree, according to the available actions.
- 3. *Simulation*: A simulation is run from the new node(s), according to the default policy, to produce an outcome.
- 4. *Backpropagation*: The simulation result is backpropagated through the selected nodes to update their statistics.

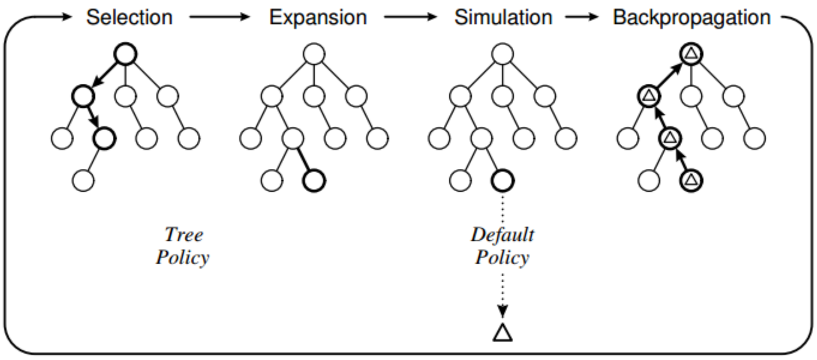

Figure 4.2: Monte Carlo Tree Search general approach

MCTS is a versatile algorithm that can be used in a variety of different settings for decision-making purposes. Specifically, there are implementations of MCTS for a lot of different multiagent games with some of them being: Settlers of Catan [24], Diplomacy [25], KriegSpiel [26] and many others. MCTS has also been used in PowerTAC in the works of SPOT agent [3], in order to select suitable strategies for the Wholesale Market.

4.2 An Interesting Equilibrium Strategy for Repeated Multiagent Zero-Sum Game Settings

The Lemonade Stand Game (LSG) is a game theoretic setting with important real-life applications. Specifically, it is a game that can provide important intuitions regarding the choices facing online advertisers, regarding which spot to bid for when participating in real-time online auctions for slots showing up in sponsored search results. In its simplest form, the lemonade game involves N lemonade vendors choosing a location to place their counter at, on the perimeter of a circular island. The utility of each vendor is determined by the distance between her, the neighbor vendors and the defined space boundaries, as depicted in Figure 4.3. In

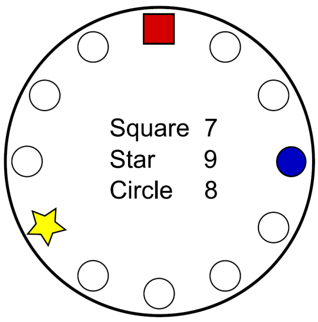

Figure 4.3: An example of Lemonade Stand Game [2]

2010 the first Lemonade Stand Game Tournament involving artificial intelligent agents took place, sponsored by Yahoo! Research, and the strategy of the winning team was shown to be the equilibrium strategy of the LSG [2]. In short, the strategy demands that one should always sit opposite of an opponent, with the purpose of maximizing their agent's utility which is translated into profit.

In general, a lesson learned from this equilibrium strategy is that in such settings we should seek to always, at each iteration, claim a large enough slice of the pie available. This strategy will ensure that any other player will be always getting lower payoffs than ourselves. In our setting, we are inspired by this equilibrium strategy and develop a strategy for the retail market that seeks to control a high portion of the market share by subscribing a large number of consumers to our services, but also restrain our "greediness" to avoid suffering huge penalties due to transmission capacity fees.

4.3 The Retail Market Module

The part we focused our research and work on was the Retail Market Module due to its extreme importance. This component's main responsibility is to publish and revoke tariffs in a way that would be profitable for the agent. Publishing and revoking tariffs alone might sound simple, but there are many aspects of the game that have to be considered before even taking any of these actions. In the following subsections, all these different aspects of the TUC-TAC agent are described in detail.

4.3.1 Preferred Tariff Types

A PowerTAC game has a specified amount of different types of power consumers, thus some distinct types of tariffs should be offered. The four distinct tariff categories are Consumption, Storage, Interruptible Consumption, and Production. However, there are more specific power consumers types for each category as seen in Figure 4.4.

So, an agent can publish tariffs for either a whole category like Storage or a specific subcategory like Thermal Storage Consumption. Also, when a "whole category" Tariff is published, it can be used by all the other subcategories as seen in Figure 4.4. For instance, when an agent publishes a Consumption tariff, along with simple consumption consumers both electric vehicles and thermal storage consumption customers can evaluate and consider subscribing to this tariff.

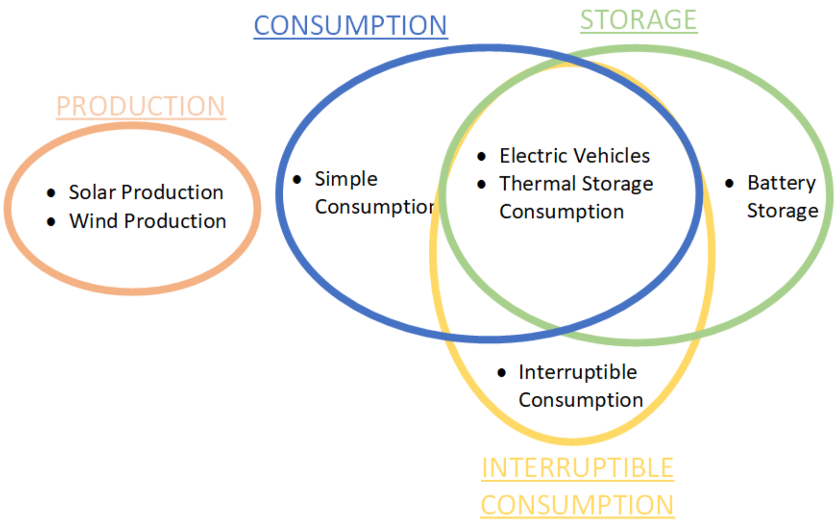

Figure 4.4: Tariff Types

In TUC-TAC's case, strategies for only 4 different tariff types are implemented. These tariffs are about *Consumption*, *Thermal Storage Consumption*, *Solar Production* and *Wind Production* costumers. After many simulations locally and online, it was found that Interruptible consumption and the other two Storage tariff types required a whole different approach to be profitable, so these were not implemented in the current TUC-TAC agent. Another reason that strategies for these specific tariff types were not further developed, was the difference in the significance of these with the simple Consumption tariff type. More specifically, the profits from the simple Consumption costumers in a small period of timeslots could be in the tens of thousands, while the profits of the other tariff types were usually at a maximum of a few thousand.

To summarize, simple Consumption tariffs were selected to be implemented and offered by our agent, because they provided TUC-TAC with an amount of profit that was in expectation significantly higher than that of other tariff types. Also, the two different sustainable Production tariffs were selected, not for their potential of making a profit, but for their ability to reduce the transmission capacity fees.

This will be further explained later. Finally, Thermal storage consumption tariffs were selected because they provided a considerable stable income. Specifically, the income from these tariffs were a few thousand from the customers and a few thousand from the balancing market. Moreover, these tariffs were considered, because it was necessary to prohibit TUC-TAC's competitors from taking the advantage of these non-Consumption tariffs.

4.3.2 Tariff Parameters

In Figure 4.6 below we can see, the main tariff parameters that TUC-TAC used.

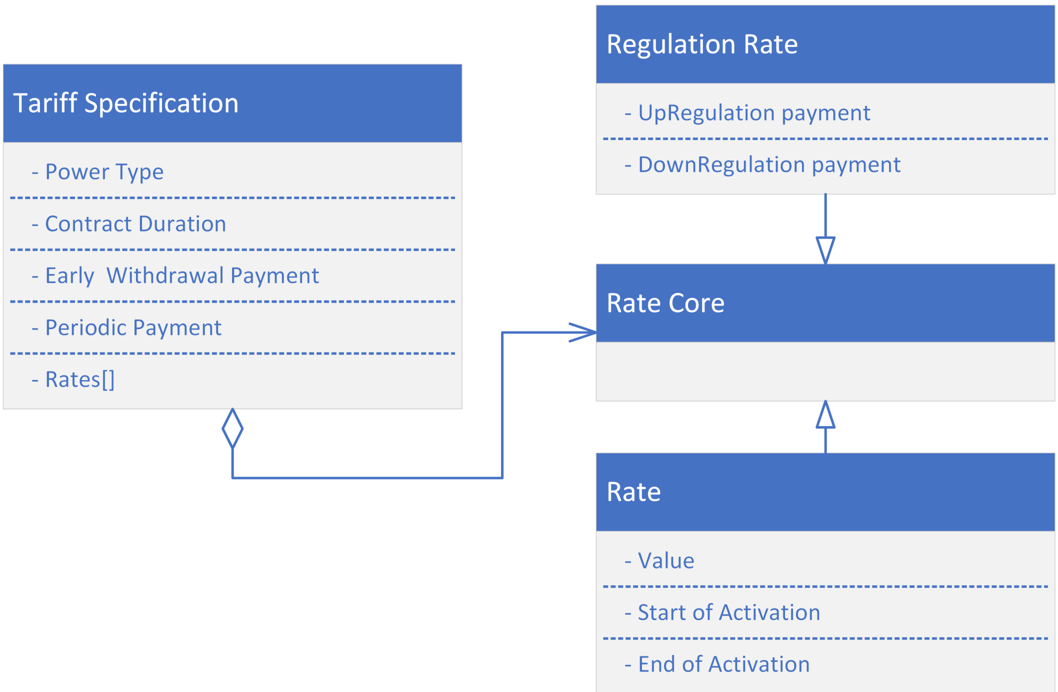

Figure 4.5: Basic Tariff Parameters used by TUC-TAC

Specifically, Power Type refers to the type of tariff. Contract duration determines the timeslots in which customers will be charged with early withdrawal fees if they decide to change their active tariff. Also, Periodic Payment as its name suggests is the money that a broker will pay periodically so it can remain subscribed to that tariff. Finally, a tariff must contain at least one Rate. Every Rate describes the payment per KWh used during a specified time of a week. For example, one Rate can refer to the payment for each day after midnight till morning and another Rate with a different price can refer to the payment for each day after morning until midnight. Regulation Rates are more simple, these are used by Storage devices to specify separate payments for use of the device for up-regulation or down-regulation.

TUC-TAC uses 48 different rates, 24 to specify the payment for each hour of weekends and 24 for each hour of the weekdays. This number was inspired by the previous work of Serkan Özdemir [19]. His idea was to try to shift the consumers' energy usage by adjusting the prices for each different timeslot in order to reduce the net demand peaks. By using that time-of-use price scheme, he claims that the peak-demand charges are significantly reduced. This technique required a net demand predictor to work. Unfortunately, in our case, there was no properly functioning predictor by the time of the finals, so though the number of the 48 rates remained and is mostly symbolic. Another reason the number 48 remained had to do with the hidden customer evaluation models. Every customer in the game indeed acts as an autonomous entity, so each one evaluates the available tariffs using some specific protocols along with some sort of randomness. After observing many games during the trials, we realized that tariffs with more than one rate were usually more attractive than tariffs with only one rate. Also, most of the other agents used a similar scheme, with more than one rate. Some of the other tariff rate numbers were 1, 12, 25, and 168. So we did not change the number of rates, even though the values of our rates were mostly random (within very strict bounds). More specifically, after obtaining the rates with the help of a completely random net demand "predictor", the rates were normalized around a target value and then repeatedly scaled-down till the standard deviation reached a particular number(0.02).

4.3.3 Objective value of a Tariff

One of the main problems a PowerTAC agent has to solve, is the evaluation of its opponents' tariffs, with the purpose to offer better ones. The difficulty of this problem derives from the complexity of the customers' evaluation model itself. Also as it was presented earlier, a tariff has many parameters to consider while evaluating its objective value. For example some of these parameters are periodic payment, rates, early withdrawal penalties, sign up bonuses and so on.

During the preparation time, we developed three different evaluation functions, while only one of them worked as intended. But even before that, a basic principle was set. Opponents' tariffs which had unusual features were considered as baits and were not evaluated. Some of these features could be very high early withdrawal penalties, unusually high periodic payments, or values of rates. Thereafter, the average value of rates was calculated using three different methods.

The first method tries to find the average value of the rates with the help of the weights which were produced from the time-of-use-technique [19]. The second method calculates the average directly by using the values of the rates without any normalization. The third method calculates the average after normalizing the values of every rate in the tariff.

The first method was quickly rejected due to the lack of a proper net demand predictor, so the other two methods remained to be considered. The second method was the one that was initially developed so TUC-TAC's parameters were fine-tuned according to the output of that method. Later, when the third method was developed and tested, the results were much worse than these of the second method, so this was rejected too. Even though the second method produced better results, we believe that the method which normalizes the rates has the potential to be superior if the agent is fine-tuned using that method. On the other hand, as far as the agent was winning there was no need to change any of these core functions. However, this is potential future work.

Finally, after finding the average value of the rates, the other tariff parameters should be taken into account too. Specifically, after discarding the "bait" tariffs the only parameter we considered important was that of the periodic payment. So, we come up with Formula 4.1 while experimenting to try to find the objective value of a tariff as it was evaluated by the customers:

$$\mathbf{ObjectiveValue} = \mathbf{Average} + \mathbf{PeriodicPayment}/\mathbf{20} - \mathbf{0.015} \tag{4.1}$$

This empirical, heuristic formula was quite accurate most of the time, and was the key factor that allowed TUC-TAC to offer more attractive tariffs than its opponents.

4.3.4 Initially published Tariffs

An important stage of a game is its beginning. After many simulations we concluded that the best initial strategy is to try to have the most customers subscribed as early as possible. In this way, when the game proceeds and the competitors publish new, better tariffs, TUC-TAC will be the first to benefit from the "early withdrawal" penalties that customers will have to pay. For that reason, TUC-TAC initially publishes consumption tariffs with rates close to its lower bound, in order to make them appealing to the customers. At the same time, it also publishes tariffs for the other three power types, but without any intention to make these especially attractive. The main rationale for this, is to make its opponents offer cheaper tariffs for customers, so they could not benefit that much. Also as mentioned earlier, these tariffs do not have the potential to bring a lot of profits, so just canceling out the competitors is enough for TUC-TAC's current strategy.

4.3.5 Main Tariff Strategy

Since the basics of the game and some "peripheral" strategy aspects have been explained, we now turn to describe in detail the strategy which was responsible for TUC-TAC's success. As mentioned earlier, the basic principle that we applied has certain analogies to the equilibrium strategy used by the winning agent of the 2010 Lemonade Stand Game tournament [2]. In short, their winning strategy was to try to always sit opposite of some other opponent so their agent could have high utility at all times. TUC-TAC's strategy is quite similar to that since its basic goal is to get half the available market share leaving the rest to the others. So by doing that, TUC-TAC expects to always have the highest income, while it shares all the fees with the other agents. Figure 4.6 below outlines the main components of the TUC-TAC strategy.

We now explain the flowchart components in detail. In the beginning, we publish our initial tariffs and then we wait for the assessment timeslots. When it is time for a reassessment of the market state, our agent first checks if any of its current tariffs are exceeding some specific dynamic bounds. The tariffs that are out of bounds get revoked, the others remain. Then we check the number of customers that are subscribed in the total of a tariff type. If the amount of the subscribed customers is higher than the *MIDDLE-BOUND* we instantly revoke our cheapest tariff and create a new one with the purpose to share the customers with the other brokers. If the amount of the subscribed customers is not higher than the *MIDDLE-BOUND* we then check the Lower-Bound. The purpose of having a *Lower-Bound*, is to remain competitive throughout a game, so, if the amount of the subscribed customers is lower than the *Lower-Bound*, we try to create and publish a tariff that is more attractive than that of our opponents. And then all

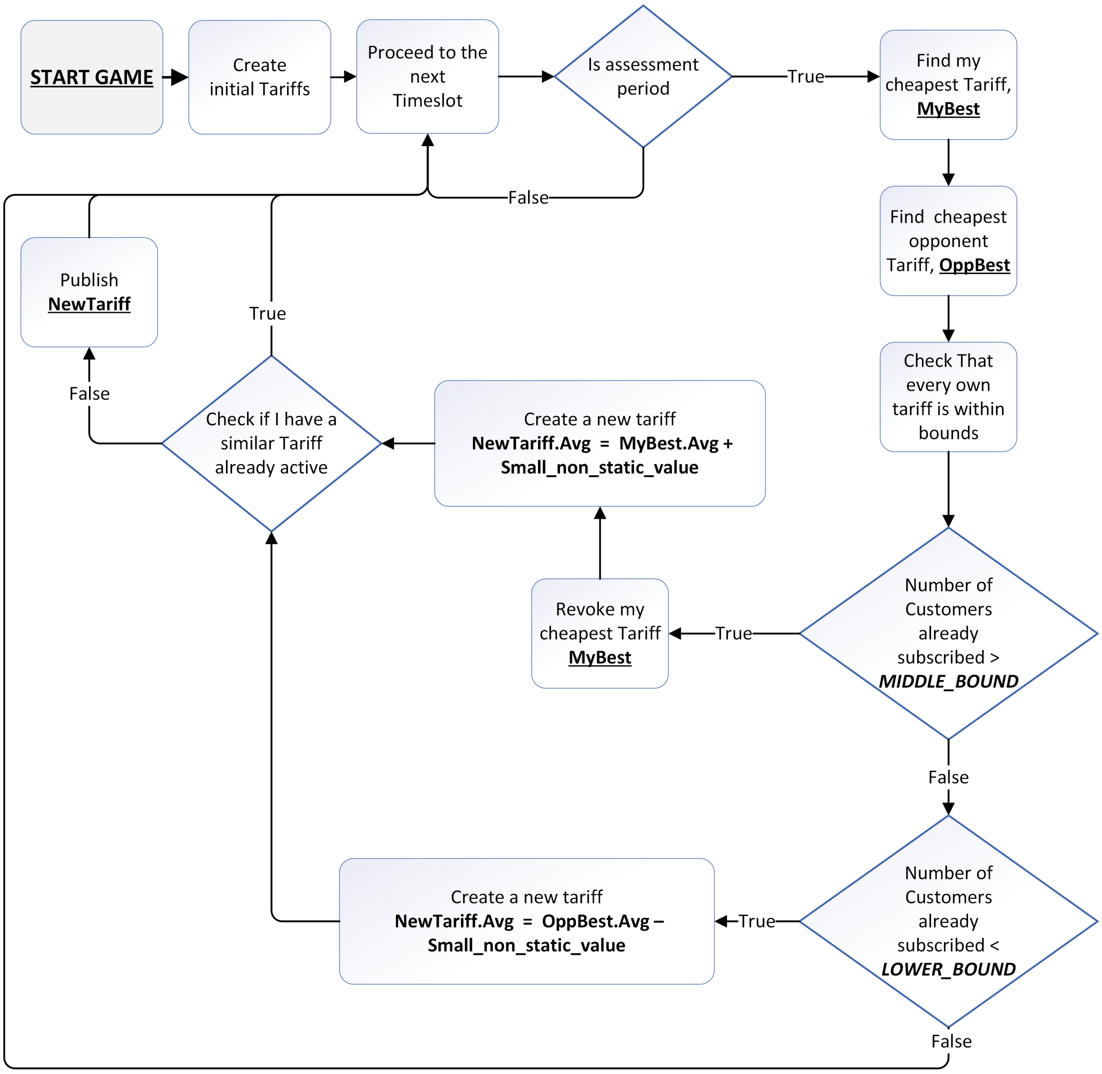

Figure 4.6: Main Consumption Tariff strategy flowchart

over again till the game ends.

As it was mentioned earlier, TUC-TAC offers 4 different types of tariffs. Also, in PowerTAC publishing and revoking tariffs is not free. So, in order to minimize the losses from publications and revokes, only the Consumption Tariffs used the main strategy because of their special significance. The rest of the tariff types use the secondary tariff strategy, which will be explained later.

There is one more basic state of the agent that is not visible in the flowchart.

That state is called "Low Customer Percentage" and as the name suggests TUC-TAC enters that state when it fails to subscribe more than a specified percentage of customers for a prolonged period of time. This is a very rare case and it occurs mostly in 3 player games when facing specific opponents. So, in order to escape from that situation, TUC-TAC lowers its tariff's bounds to offer more attractive tariffs than its opponents.

Besides the "Low customer" State there also exists an opposite one. This is a special case that triggers when more than 90% of the customers are subscribed in the total of TUC-TAC's Consumption Tariffs. This is a very dangerous situation because when a broker has the monopoly, it might get all the profit but it will also have to cover all of the Transmission capacity fees, which is disastrous. So, when that happens TUC-TAC immediately revokes all tariffs that are below a specified bound keeping only a portion of the customers.

4.3.6 Secondary Tariff Strategy

Besides Consumption Tariffs that used the previous strategy, the other three tariffs offered by TUC-TAC used the secondary tariff strategy. This strategy follows the same principles with the main one, but in this case the assessment period is 168 timeslots so the publication and revoking fees are greatly reduced. Also, it is not necessary to constantly control the market because of the low potential to make profits with these tariffs. So TUC-TAC's strategy aimed at first to increase the market prices of these three types of tariffs and then to benefit directly from them. Furthermore, it was not considered crucial to compete for these tariff types in the 8-player games, thus only some initial, "safe", ones were published. After seeing the results of 8-player games we now believe that this decision was most probably wrong. Regardless, Figure 4.7 depicts the secondary tariff strategy.

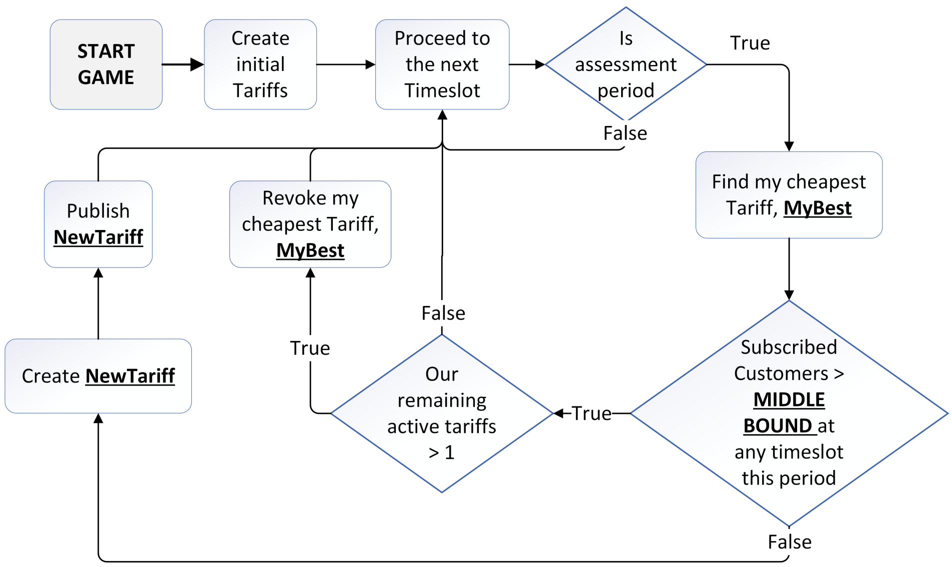

Figure 4.7: Secondary Tariff strategy flowchart

4.3.7 Agent Parameters and Bounds

PowerTAC is a complex environment so finding the optimal strategy is not obvious. Despite that, we come up empirically with an equation rule which, as long as it was satisfied, TUC-TAC was having a positive balance at the end of an assessment period. These assessment periods were 168 timeslots long,i.e. their length was the same as that of the periods over which the Transmission Capacity Fees were evaluating and charging. In Formula 4.2 below, $Income_{TUC-TAC}$ represents the total incomes from any sources of TUC-TAC during a timeslot period, while $TCFees_{TUC-TAC}$ are the transmission capacity fees charged at the end of the same period. On the other side, the $Income_{others}$ and $TCFees_{others}$ represent the sum of all the other brokers' income and fees respectively. The $Offset$ is a dynamic parameter that changes depending on the type of the game (game location,and number of players).

$$\mathbf{Income_{TUC-TAC} - TCFees_{TUC-TAC} + Offset > Income_{others} - TCFees_{others}} \tag{4.2}$$

So, with formula 4.2 in mind, the TUC-TAC's parameters and bounds were fine-tuned to observe that equation.

Furthermore, an additional observation helped us fine-tune the agent. This observation was that TUC-TAC usually had to pay increased balancing fees during late December, January, and early February.This phenomenon probably occurs because of the extreme weather that was not allowing solar panels to produce energy, also the electricity needs were higher too, probably because of house heating during the winter. So during these months, the market share bounds were reduced by 5% to lessen the market share our agent would get. In Table 4.1 below we can see the precise values of each bound used by TUC-TAC 2020.

Bounds	Normal Case	Winter Months	End of Game
UPPER_BOUND	92.5	87.5	92.5
MIDDLE_BOUND	62.5	57.5	55
DO_NOTHING_BOUND	55	50	55
LOWER_BOUND	45	40	35
LOW_PERCENTAGE_BOUND	35	30	25

Table 4.1: Boundary values used in this year's TUC-TAC agent

All these numbers are referring to the percentage of the subscribed customers of a tariff type. Specifically, the *UPPER-BOUND* is used when TUC-TAC has subscribed almost all the available, as mentioned earlier in section 4.3.5. The *LOW-PERCENTAGE-BOUND* is used in the exact opposite situation when TUC-TAC is stuck in a none profitable state. The *MIDDLE-BOUND* and the *LOWER-BOUND* are mainly used by the Main Tariff Strategy as depicted in Figure 4.6. At last, the *DO-NOTHING-BOUND* determines a state in which there is no need for any action from TUC-TAC.

In that Table, we can also see the "End of Game" parameter. All the timeslots with numbers greater than 1500 are considered to correspond to "End of Game" ones. The idea behind that is that at the end of a game we do not want to take any unnecessary risk, so to achieve that we aim to decrease our agent's market by lowering the "Lower" and the "Middle" Bound by 10%.

4.4 The Wholesale Market Module

The second but equally important module of our agent is the Wholesale Market one. Its main responsibilities are to buy and sell energy in the double auction of the wholesale market. In order to be effective though, it requires finding the best bids so the customers would not have to resort to Balancing Utility to get their energy. So with a few words, if this module fails to acquire the energy required by the subscribed customers, the Balancing utility will charge higher every single KWh that was not reserved by TUC-TAC, thus creating many unnecessary penalties.

4.4.1 Monte Carlo Tree Search in TUC-TAC 2020

The main algorithm implemented in this module was a variation of the Monte Carlo Tree search method previously developed by Chowdbury et Al.(2018) [3]. In PowerTAC's case, the double auction of the wholesale market is a complex action-space which requires fast and precise actions in order to be profitable. So, MCTS was selected for its ability to rapidly traverse through huge decision trees and find the best action.

Specifically, in Algorithm 1 we can see the pseudo-code of our bidding strategy. At first, the root node is created along with its children. Each child represents a different *bid* in the double auction, and the values for each one are generated by combining the predicted limit price and a standard observed deviation. In our current implementation, the limit price predictor is just a random number generator enhanced with some empirically added bounds. After that, the MCTS algorithm continues to traverse the tree by expanding the nodes and simulating some auction outcomes. After every iteration all nodes that were part of the current path are updated, specifically, we save information about visit count and average unit cost so we can calculate the UCT value for each node when needed. In conclusion, we can infer that the concept of this algorithm is indeed suitable for this setting, and can be especially useful, judging by the results of Chowdhury et al [3], but the lack of a proper predictor in our case makes our current wholesale market approach completely naive. For this reason, we are already working towards creating a limit price predictor for the TUC-TAC 2021.

Algorithm 1 Calculate Bids for wholesale Market using MCTS (timeslot t)

```
energyToBuy = neededMWH(t)
for i < NUMBER_OF_ITERATIONS do
    curNode = root
    curNode.GenerateKids()
    while energyToBuy > 0 do
        if curNode.hasUnexploredKids() then
            curNode = GetRandomUnvisitedChild()
            while energyToBuy > 0 do
                limitPrice = computeLimmitPrice(t + timeslots_ahead)
                clearingPrice = GetRandomGaussianNumber()
                if limitPrice > clearingPrice then
                    Csim = energyToBuy * clearingPrice
                end if
            end while
            break
        else
            curNode = GetBestUCTChild()
            energyToBuy = Simulate(curNode)
        end if
    end while
    Cavg = Csim/energyToBuy
    Bacpropagate(Cavg)
end for
bid = GetBestRootChild().bid
BidInAuction(bid, t)
```

4.5 Monitoring Module

The last component which completes TUC-TAC-2020 is the monitoring module. This module's main responsibilities are to keep track of the game and produce logs suitable for the process of the after-game "fine-tuning" of the agent. So this component has very similar functions to that of a logger.

Specifically, this module processes and keeps track of information about :

- Bootstrap data in order to attain information about the costumers
- Total Energy Consumption and Production for every timeslot
- The profits/losses of all the tariffs offered by TUC-TAC
- Distribution, Balancing and Transmission Capacity fees
- Incomes and losses from sold/bought Energy in the Wholesale Market
- Gains from the imbalances occurred by Thermal Storage Consumption tariffs
- Total profits in general and Total profits from each tariff type

It also stores data about:

- Definitions of TUC-TAC's active tariffs and customer counts for each one
- Definitions of Opponents' tariffs

At the end of every PowerTAC game, a log file is produced containing all this information. This file was employed to fine-tune the agent's parameters empirically, via human examination and actions. In addition to that file, another one is produced. The second one is a spreadsheet file (.xlsx) which contains information about the energy usages for each timeslot and it was used to evaluate the performance of the various implementations of the net demand predictors.

Chapter 5

Experiments and Results

In this chapter we are going to present the experiments and results related to this thesis. Since this is a thesis describing a competitive agent that participated in an international competition (and won it), this chapter is structured as follows: we begin by describing first steps towards our agent's implementation, the results of the 2020 PowerTAC trials and qualifiers in which TUC-TAC participated, and the lessons learned and strategy modifications performed as a result. We then proceed to present the results of the PowerTAC 2020 finals, in which TUC-TAC prevailed; and finally, we conduct an extensive post-tournament analysis, in order to draw important lessons from this endeavor.

5.1 Preparation and Early Development

The 2020 competition was the first time a TUC team took part in PowerTAC. So, in order to catch up with the current state of the art implementations, we carefully studied the literature relating to past agents' strategies. In the end, the ones we focused more and employed ideas from the papers by Serkan et al (2017) [19], the paper from Chodhury et al (2018) [3] and the paper of the VidyutVanika agent from Gosh et al (2019) [18].

The first paper of (Serkan et al,2017) describes a retail market strategy where a variance of a genetic algorithm is used to find the best tariffs in the market. The main reason we selected this implementation as a starting point for our work, was because AgentUDE used it and finished first in the 2017 PowerTAC competition. We also started experimenting with the implementation of Chodhury et al (2018) which was a Monte Carlo tree search variance for the bidding in the periodic double auction of the wholesale market. Even though the SPOT agent which used this implementation did not have a very good overall performance in the previous competitions, we selected their techniques because we recognized that their wholesale market performance was one of the best currently implemented for PowerTAC. The last paper from the VidyutVanika agent was not used directly to implement a technique, but we saw it as an example of how a successful agent was structured (VidyutVanika finished second in the 2018 competition).

Each year, before the official qualifiers and finals take place, a series of 2-4 trial competitions are scheduled by the PowerTAC organizers. This is a great opportunity for the participants to test new implementations. So after deciding which technologies to use in our agent, we got back to work having the first trial as a target to test our progress. Although the first two qualifiers(April and July 2020) were scheduled, our agent was not ready to compete with others at that time. Later in the August trials, we felt that TUC-TAC was ready to compete with others and we were eager to see its performance. Unfortunately, even after a few games in the trials, we realized that there was much room for improvement regarding our implementation. In Table 5.1 below we can see the normalized results from this trial.

Broker	Type 6	Type 4	Type 2	Total (Normalized)
VidyutVanika	0.689	0.788	0.447	1.924
Mertacor2020	0.561	0.845	0.447	1.853
SPOT	0.566	0.240	0.447	1.253
xameleon	0.394	0.149	0.447	0.990
EWIIS3_2020	-0.036	0.118	0.447	0.529
TUC_TAC	-2.175	-2.138	-2.236	-6.550

Table 5.1: August's Trial Results

In these trials, our agent was always trying to acquire all of the market shares. It was obvious then, that the lack of boundaries in the tariff generation was piling up all of the transmission capacity fees on our agent, resulting in these huge negative scores. It did not take long to understand that the genetic algorithm itself was not enough to handle the huge burden of the capacity fees. So after some team meetings and a lot of tests, we created the following strategy which is depicted in Figure 5.1.

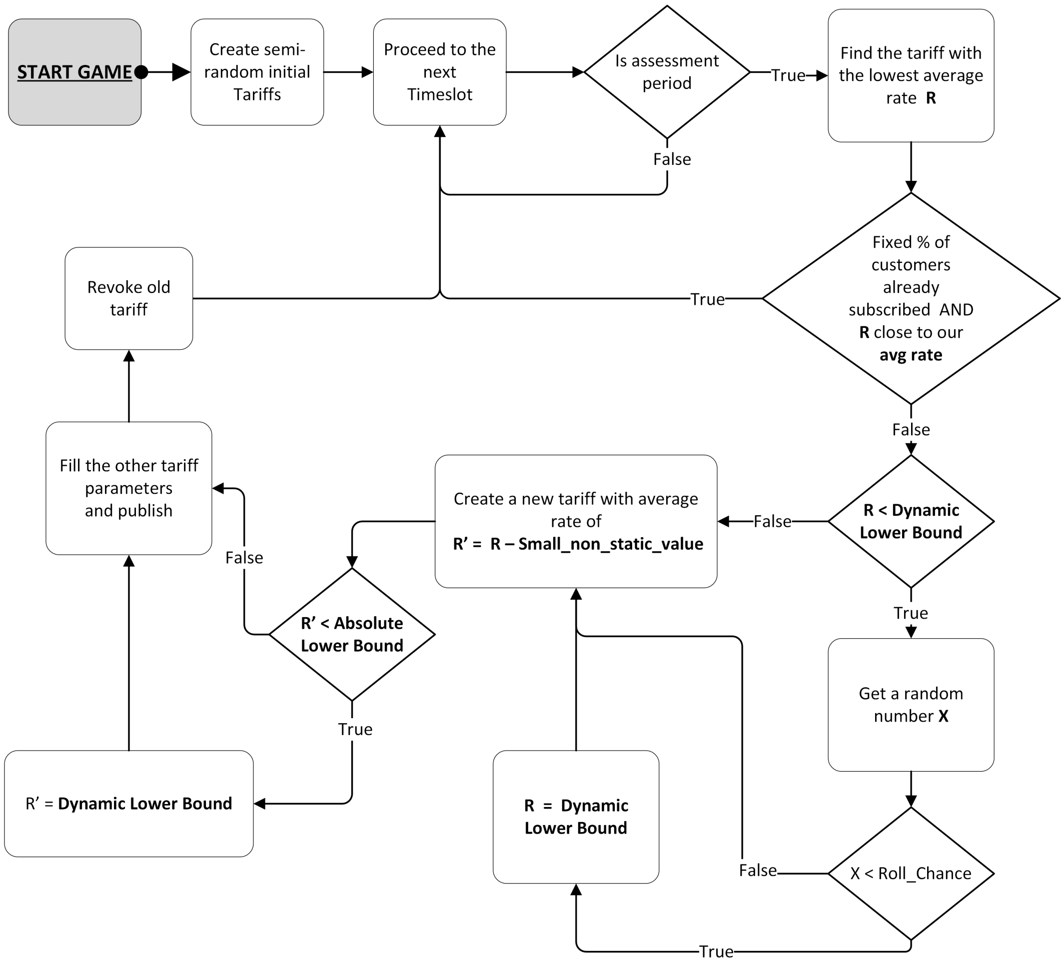

Figure 5.1: First version of the TUC-TAC retail market strategy used in September's 2020 Trial

This strategy put some distinct boundaries to the tariff creation while keeping the random aspect of the genetic algorithm. Despite us being only cautiously optimistic after these modifications, this strategy worked especially well in the final trial of September. The results of that trial can be seen in Table 5.2 below. It is clear that our total score was far better than that of August's. After processing the logs of the games we found out that the main reason TUC-TAC had these huge losses was the wholesale market. In detail, as it was mentioned earlier, when a broker cannot fulfill its customer's energy demands the Balancing Utility charges that broker for the KWh it did not reserve at a much higher price.

Broker	Type 5	Type 3	Type 2	Total (Normalized)
Mertacor2020	0.749	1.105	1.699	3.553
EWIIS3_2020	1.195	0.738	0.038	1.971
VidyutVanika	-0.134	0.468	-0.432	-0.098
CrocodileAgent2020	-0.082	-0.741	0.081	-0.742
TUC_TAC	-1.728	-1.570	-1.385	-4.683

Table 5.2: September's Trial Results

So after implementing a dominating retail market strategy, we focused on improving the smaller details of it as well as creating a better wholesale market strategy. To further improve the retail market strategy, some features had to be added to it. The first feature that was added was the support of multiple active tariffs. By using multiple tariffs of a specific type the agent benefits too from the early-withdrawal fees parameter thus increasing the final money of the broker. Moreover, for our agent to have better control of the market state more special cases had to be added. The exact flow chart of that strategy is described as the "Main tariff Strategy" in Figure 4.6.

At the same time, our team was working towards the implementation of a net demand predictor and a wholesale market prices predictor. The first would have been used to decrease the transmission capacity fees and the second was essential for the strategy of the wholesale market. Unfortunately, no appropriately functioning version of these was ever ready in time. So we had to improvise and experimentally create a substitute for these. Luckily the changes that were made in the retail market module were enough. On the other hand, the changes made in the wholesale market module were neither profitable nor harmful for TUC-TAC, but these were all we could do at that point.

Finally our agent competed in the Qualifier round of October and dominated over all other brokers. The results of the qualifiers can be seen in Table 5.3 below.

Broker	Type 8	Type 5	Type 3	Total (Normalized)
TUC_TAC	1.745	1.456	1.823	5.025
Mertacor2020	0.467	1.513	1.024	3.004
EWIIS3_2020	0.892	0.341	0.080	1.153
CrocodileAgent2020	-0.258	-0.243	0.418	-0.083
VidyutVanika	0.161	0.045	-0.812	-0.607
SPOT	-0.598	-0.715	-0.987	-2.300
xameleon	-1.746	-1.087	-0.040	-2.874
COLDPOWER2020	-0.663	-1.311	-1.345	-3.319

Table 5.3: October's Qualifier Results

After these results, we were optimistic that our work paid off, but we did not rest yet because the finals were near. In order to further prepare our agent for the finals, some other aspects of the agent were added. One of them was the secondary tariff strategy (Figure 4.7) which helped in decreasing the capacity fees while it canceled our opponents' strategies that were based on production tariffs. The other less significant addition was that of the "Low-Percentage" state which helped in getting out of some not profitable situations.

In brief, that was the preparation and the early development we made for the PowerTAC 2020 finals.

5.2 PowerTAC 2020 participation results

The 2020 PowerTAC competition consisted of 8 teams from around the globe (USA, Germany, Mexico, Croatia, India, and Greece) competing for first place in the competition. Specifically, each agent participated in 40 eight-player games, 105 five-player, and 63 three-player games. This year's competition had some very interesting games judging by the fact that each game type had a different score leader. More details about the scoreboard can be seen below in Table 5.4. So after this tough competition, it would be wise to analyze each different game type and see which agent was performing better in each scenario.

Broker	**Type 8**	**Type 5**	**Type 3**	**Total (Normalized)**
TUC_TAC	0.494	2.123	0.991	3.607
Mertacor2020	0.546	-0.196	1.576	1.926
CrocodileAgent2020	0.207	-0.142	1.105	1.170
EWIIS3_2020	0.684	0.158	-0.215	0.627
VidyutVanika	0.282	0.395	-0.531	0.145
SPOT	0.252	-0.127	-0.972	-0.847
COLDPOWER2020	0.141	-0.474	-1.235	-1.568
xameleon	-2.605	-1.736	-0.718	-5.059

Table 5.4: November's Final Normalized Results

The exact aggregate scores of all games for each different type of PowerTAC 2020 competition are depicted in Figure 5.5. Specifically, the un-normalized, actual scores can give us more information about how each agent actually performed in different game types. On the other hand, normalized[1] scores give us a better overall picture of which agent is better in each game type.

[1]An agents normalized score denotes the fraction of its actual score over the average score across all agents.

Broker	Type 8	Type 5	Type 3	Total
TUC_TAC	2,962,734	121,293,603	131,705,373	255,961,710
Mertacor2020	3,726,679	5,491,646	164,138,372	173,356,697
CrocodileAgent2020	-1,250,817	8,178,165	138,004,184	144,931,532
EWIIS3_2020	5,760,600	23,177,825	64,875,477	93,813,901
VidyutVanika	-152,168	34,985,883	47,364,189	82,197,904
SPOT	-597,434	8,945,948	22,952,988	31,301,502
COLDPOWER2020	-2,233,261	-8,372,169	8,379,713	-2,225,717
xameleon	-42,622,991	-71,444,113	37,005,980	-77,061,123

Table 5.5: November's Final not-normalized Results

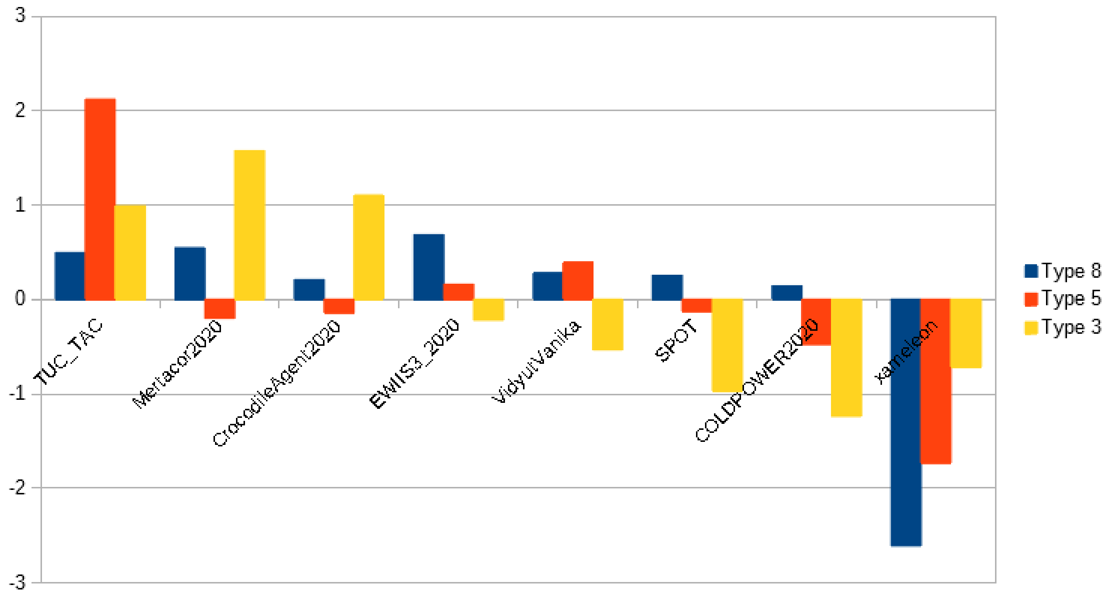

Figure 5.2: Final Normalized Scores of PowerTAC 2020

At first, we can see in Figure 5.3 below, the average score of all games played. The vertical axis will always show the score while the horizontal axis presents the name of the broker in each of the three different scenarios. These scenarios are games with three-players, five-players, and eight-players.

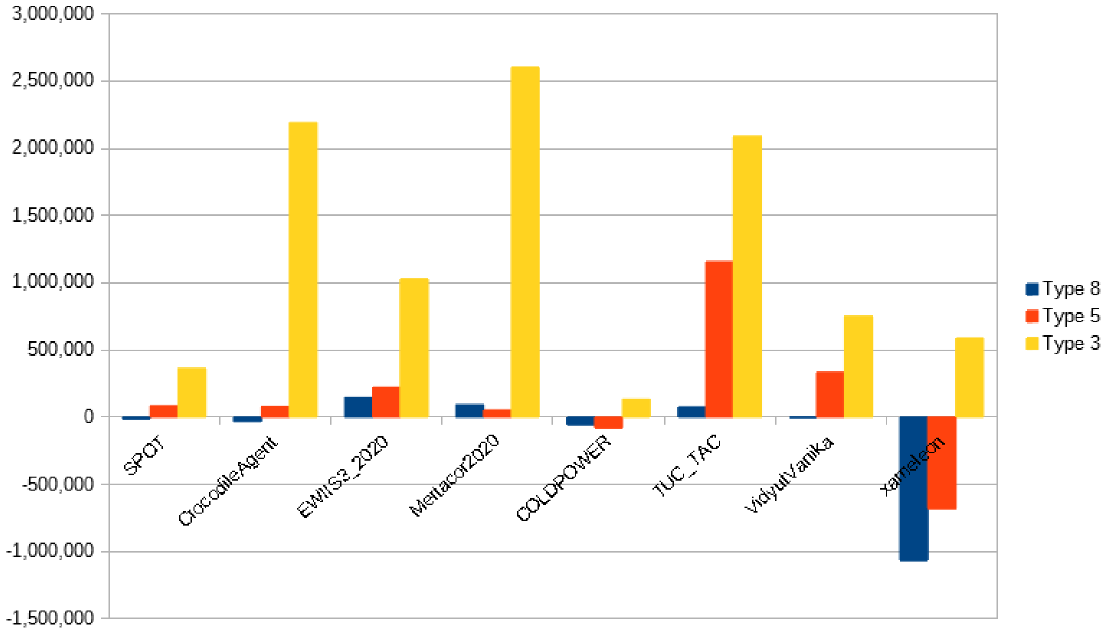

Figure 5.3: Average scores of all PowerTAC 2020 games

5.3 PowerTAC 2020 post-tournament analysis

In PowerTAC 2020 each participant had a different game strategy, even though most of them used similar technologies like Markov Decision Processes, Q-Learning, or simpler approaches such as us for retail market decision making. For example, TUC-TAC and Mertacor2020 were quite aggressive in the retail market regardless of the number of players in the game, specifically, Mertacor used an aggressive decision-making strategy informed by offline reinforcement learning. But in the end, TUC-TAC's smaller details such as faster response times, and the offering of more attractive tariffs, allowed it to have an advantage over Mertacor, and this is why TUC-TAC won most of the games it participated in. At the same time, CrocodileAgent2020 was especially aggressive in 3 player games but was too afraid to do the same in more-than-three player categories, as we will see in the rest of this analysis. Moreover, ColdPower, Spot, VidyutVanika, and EWIIS3_2020 had a similar "conservative" behavior in the retail market judging by their lower average scores in most games. It is also important to note here that apart from the retail market strategy failure of Spot and VidyutVanika, both of these agents were the ones that were performing the best in the wholesale market because they managed to generate profit by selling and buying energy there, something that will not be further investigated in this thesis. On the other hand, Xameleon implemented a greedy strategy that did not perform well, probably because of the high fees it had to pay and some flaws in its design as it was later mentioned by its developer. In the rest of this chapter, we will provide an extended post-tournament analysis with detailed Figures.

5.3.1 Categorization by balancing fees

Balancing fees are the fees that are applied to the agents by the Balancing market when they fail to procure the required energy. The most common reason a broker might fail to accumulate the required energy, by its customers, is very high wholesale market prices.

We can see from the Figure5.3, that the average score of most brokers in type 8 games is similar. Most specifically, the standard deviation between the first seven agents is quite small (74,000), in contrast with Xameleon which had some especially high losses in type 8 games. The standard deviation of that category along with Xameleon is 393,000.

However, for the tier 5 games, the performance of TUC-TAC was outstanding managing to have almost three times greater average score than the second contestant. In more details, the average score of TUC-TAC was 1,155,000, while VidyutVanika which was the second, had only an average score of 333,000. The standard deviation of type 5 games is 508,514.

Unlike the two previous game types, the tier 3 games were the most competitive, having agent Mertacor as the leader followed by CrocodileAgent and TUC-TAC closely. As we can see in Figure 5.3, the standard deviation of 3-player games is higher than any other game type. In particular, the standard deviation is 940,000.

But how did these final scores occur? In order to understand that, we are going to break down the games and see how each broker was performing in different game types besides the classic categories.

From our viewpoint, there were two distinct types of games in this year's finals. The "regular games" and the "special Phoenix games". We define as "special Phoenix games" the games that have extended periods of timeslots with unusually high wholesale market prices. In such situations, agents that have not been careful to buy substantial amounts of energy early on, would have to buy energy in very high prices in the wholesale market. As it was observed in PowerTAC 2020, the leading agents were not prepared for this scenario, thus they could not obtain the required energy during these time periods, and resulting in very high balancing fees for each one of them. In our case, this phenomenon usually occurred during the summertime of games located in Phoenix. However most of the time, these games do not contain any other feature that makes them special. For example, another feature of some other games is unstable weather leading to higher than usual net demand peaks.

Our agent won the majority of games in every "classic" category, but only managed to have the best overall score in the five player games while being third in the other two game types. More details about the wins of TUC TAC in each game type (clearly depicting performance in "Phoenix" games also), can be seen in the following Table 5.6.

Below we can observe how each agent performs in different types of games. Specifically, we will depict the average scores of Regular games in Figure 5.4 first;

	Type 8	Type 5	Type 3
Regular Games	28/34	81/102	47/57
"Phoenix" Games	0/6	0/3	0/6
Total	**28/40**	**81/105**	**47/63**

Table 5.6: Total wins of TUC-TAC in the finals

then, Figure 5.5 depicts performance in the "Phoenix" games. This comparison will make clear how substantially different the rest of the Phoenix games were.

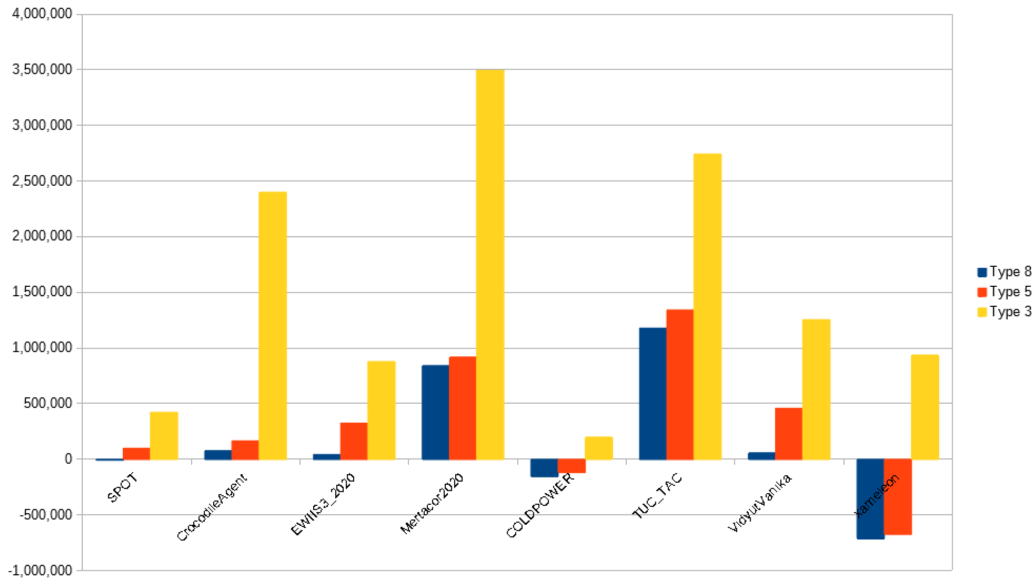

Figure 5.4: Average scores of Regular games

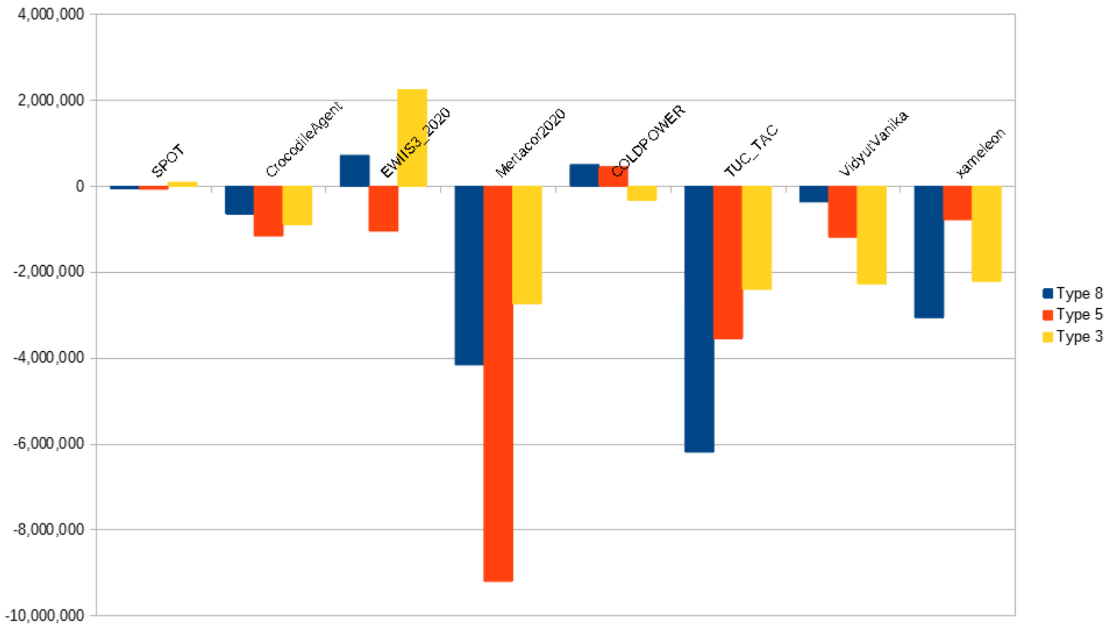

Figure 5.5: Average scores of "Phoenix" games

By comparing the two previous graphs we can clearly see how different "Phoenix" games are especially for the winner broker TUC-TAC and the runner-up Mertacor. The main reason our agent was under-performing in these games was a flaw in the design of our wholesale module, but as we can observe almost none of the other participants was prepared for these games as well. As mentioned earlier in Section 4.4, this flaw had to do with the inability of TUC-TAC to buy enough energy from the wholesale market to provide it to its customers, thus resulting in high penalties for it. However, it seems like these scenarios too were very profitable for EWIIS who had a very stable performance throughout the tournament. Even though EWIIS was very stable, it was not enough to win the tournament. That shows that in order for an agent to win a tournament some aggressive actions should be taken.

5.3.2 TUC-TAC's impact on different games

In this section we will try to show how the existence of TUC-TAC agent affected the overall outcome of the games.

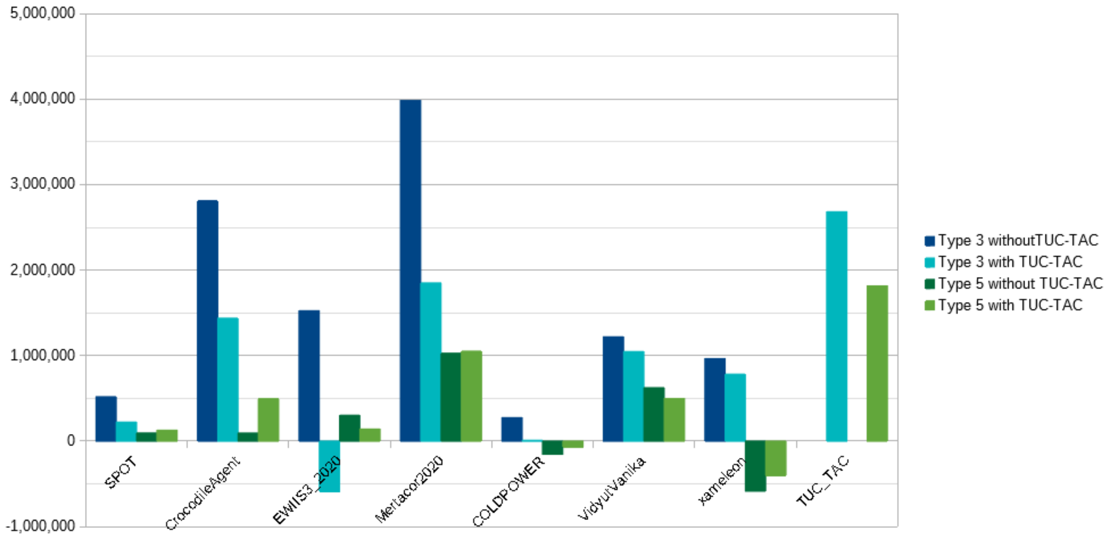

Figure 5.6: Average scores of the agents in the Regular games with and without TUC-TAC

The most notable thing about Figure 5.6 is the average score of tier 3 games. We can see that when TUC-TAC is participating in a game it tends to have the highest score. However, when TUC-TAC is not part of a game the other brokers, especially Mertacor and CrocodileAgent, manage to almost double the average score they had when TUC-TAC was participating. From this fact, we can infer two possible explanations. At first, it is obvious that when our agent is part of a game it manages to beat its opponents by limiting their profits, but when it is not part of a game essentially no one is there to limit their profits resulting in that huge score difference. At the same time, after looking at the individual scores of some games and Figures 5.7, 5.8, and 5.9 below, we can assume that the agents Mertacor and Crocodile were more successful than TUC-TAC in tier 3 games when they did not compete with each other.

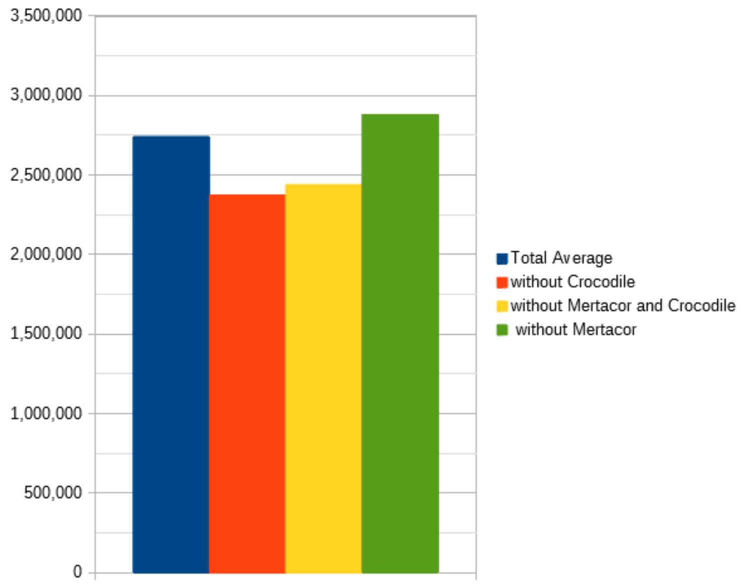

Figure 5.7: Average score of TUC-TAC for type 3 Regular games

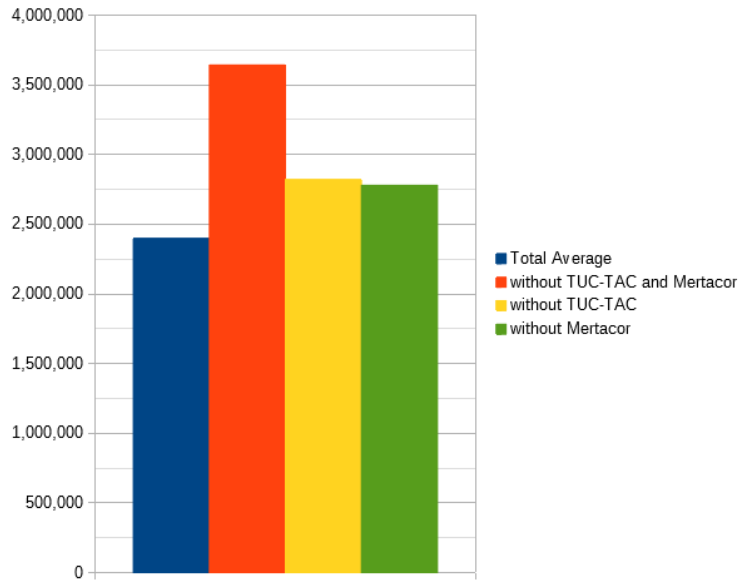

Figure 5.8: Average score of CrocodileAgent for type 3 Regular games

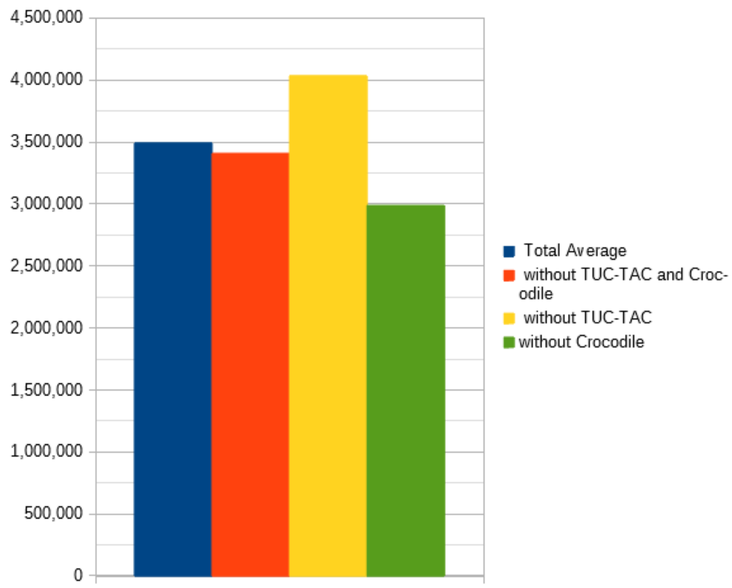

Figure 5.9: Average score of Mertacor for type 3 Regular games

By observing the previous graphs we can see how each of the three best (type 3) agents perform when their main competitors are not part of a game. The most impressive graph is that of CrocodileAgent. As it seems when both TUC-TAC and Mertacor are absent, Crocodile's average score is over 1 million higher, while when only one of the main competitors is absent its average score is almost half a million higher. After seeing that, it is fair to say that Crocodile performs better when the competition is weaker, thus that is the reason it gets second place in type 3 games.

In addition, the Mertacor agent appears to have a better performance when TUC-TAC was *not* part of a 3 player game (half a million higher than the total average), while its average score dropped by half a million when Crocodile was not part of the game. Even though, Mertacor's performance had much variances the average score in every case was still higher than every other broker in the 3-player games. Judging by that,it is not a surprise that Mertacor got first place in the 3 player games.

However, TUC-TAC's performance is not really that much affected by the absence of its main competitors. This of course is a plus in the sense the agent's performance is stable, but at the same time it signifies that TUC TAC cannot exploit weaker agents that well, unlike CrocodileAgent and Mertacor. Nevertheless, TUC-TAC's stability allowed it to get third place in the 3 player games.

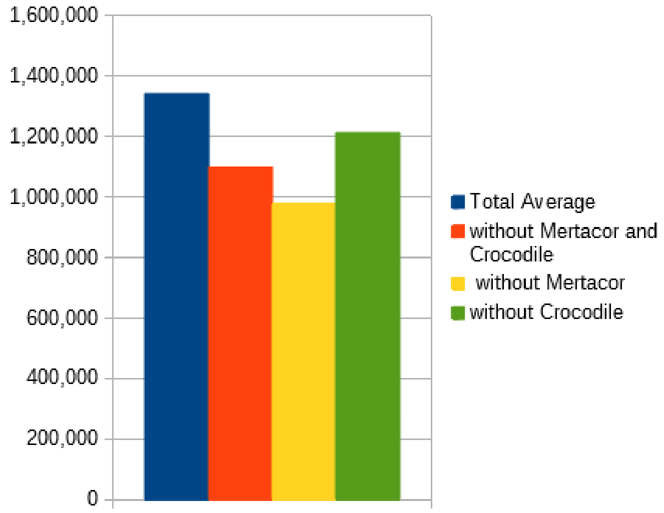

Figure 5.10: Average score of TUC-TAC for type 5 Regular games

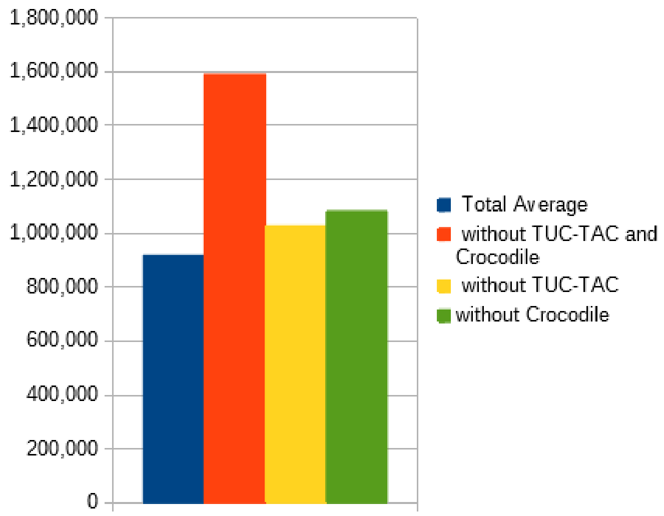

Figure 5.11: Average score of Mertacor for type 5 Regular games

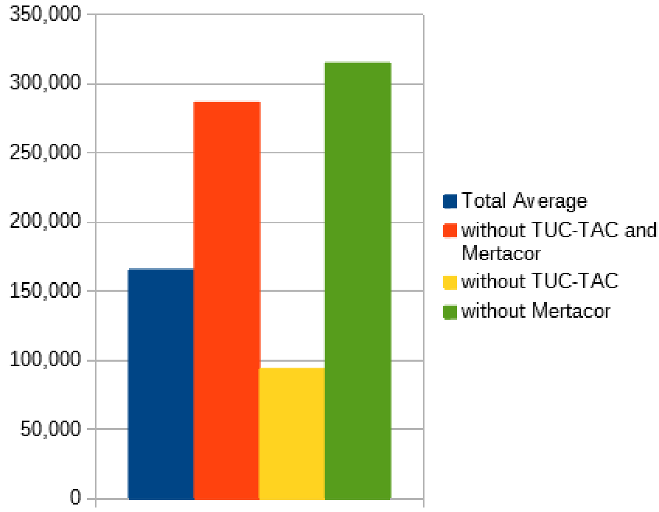

Figure 5.12: Average score of CrocodileAgent for type 5 Regular games

In previous Figures 5.10, 5.12, and 5.11 we can observe the performance of the three best agents in the Regular 5-player games when their main competitors are not part of a game.

At first, we notice that TUC-TAC's performance is quite stable in general (Figure 5.10). In addition to that, we can infer that TUC-TAC has better results when Mertacor is part of a game. This happens because the combination of these two highly competitive agents greatly reduces the market share of the other participating agents, thus resulting in higher profits for TUC-TAC and Mertacor.

In Mertacor's case (Figure 5.11), we can see that the presence of TUC-TAC and CrocodileAgent greatly reduces his average income in a 5-player game. On the other hand, this fact shows that Mertacor is better at exploiting the rest of the agents when there is no direct competition (like when TUC-TAC and Crocodile are participating).

Lastly (Figure 5.12), Crocodile seems to depend on TUC-TAC up to a point to generate profit in this game type. Specifically, we can see that Crocodile's strategy in 5-player games is not working as well as in 3-player games judging by the fact that its average score in every category (of 5-player games) is very low.

In the following Figure 5.13 we can observe how the presence of our agent assisted in balancing the losses of other agents from “Phoenix games”.

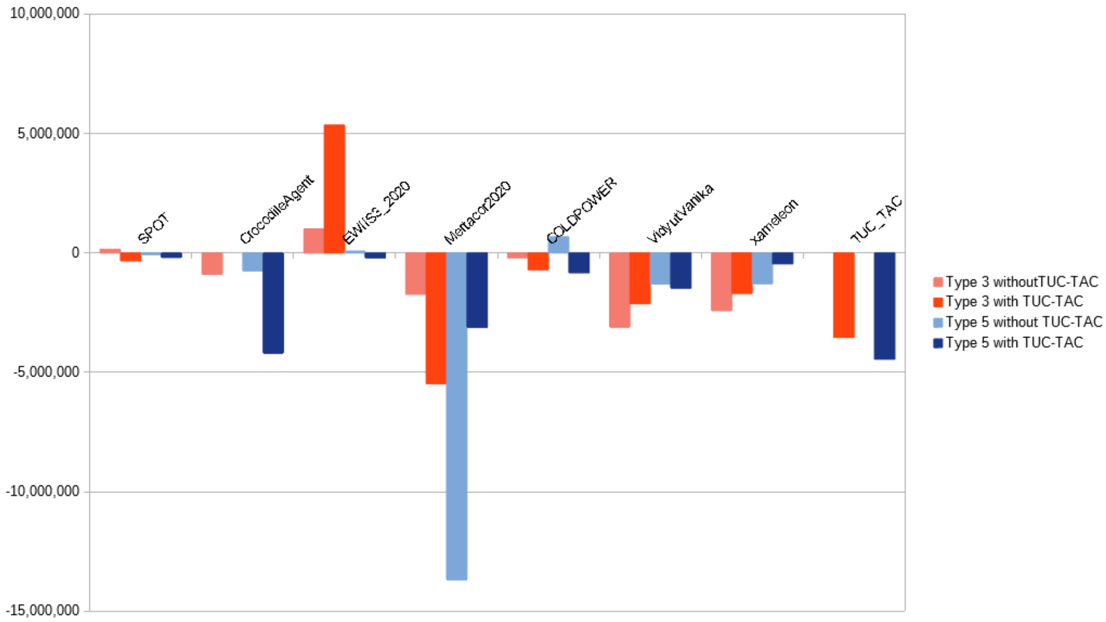

Figure 5.13: Average scores of the agents in the "Phoenix" games with and without TUC-TAC

As we can see Mertacor’s performance in tier 5 “Phoenix” games was quite devastating for itself when TUC-TAC was absent, while it was much better when TUC-TAC was part of the game. On the other hand, the EWIIS3_2020 agent was very stable in "Phoenix" games especially when TUC-TAC was part of the game, as we can clearly see in Figure 5.13. The reason behind EWIIS3_2020’s success had to do with its unique strategy. Specifically, that agent tried to maximize its profit by acquiring small market shares, but with tariffs that would benefit both the client and the broker regardless of the game state. Of course, by choosing that strategy, that agent was limited to a lower amount of profit, so this is the reason that agent finished 4th overall (Table 5.4). In conclusion, it seems that our agent’s presence was a significant factor affecting the performance of other agents, both in regular and "Phoenix" games.

5.3.3 Categorization by transmission capacity fees

Besides the main game categorization into "Regular" and "Phoenix" games we put forward earlier, in our view there exists another games categorization that is perhaps of equal importance with the former. This one has to do with the amount of total transmission capacity fees paid by the brokers throughout a game. It is well understood that if these fees were not existent, the optimal strategy in PowerTAC would be to always underbid the opponents' tariffs, aiming to acquire the whole customer base, since there would be no game mechanism to punish this aggressive behavior. Fortunately, the transmission capacity fees exist to bring balance between the possible tariff prices and the number of customers a broker can get.

So after understanding how important those fees are and how these generally affect the game in theory, we present below the results of PowerTAC 2020 for each tier sorted by the total exceeding MWh paid by the brokers as transmission capacity fees. By presenting these graphs we clearly demonstrate that our retail market strategy works well, regardless of the amount of the fees. In the following graphs, we are excluding the "special Phoenix" games because their huge losses are not associated with the capacity fees, and their inclusion would thus only result in unnecessary "noise" in our graphs.

At first we will see how the total capacity fees affect the cumulative score of 8-player games, as seen in Figure 5.14.

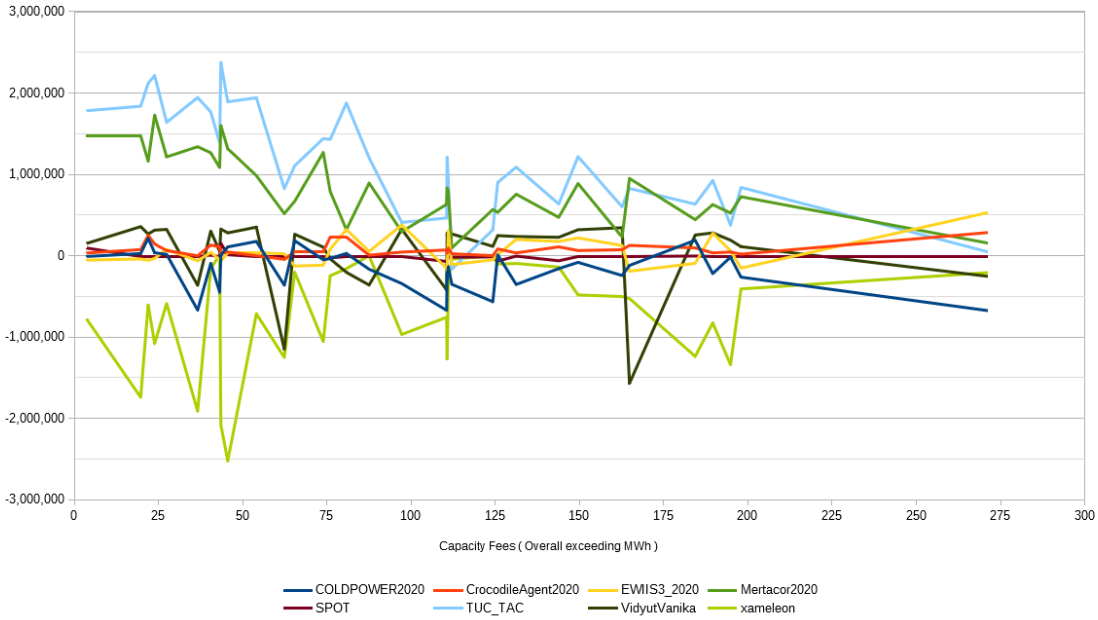

Figure 5.14: Scores of type 8 Regular games

We can clearly observe that our strategy works exceptionally well when the capacity fees are low but it tends to worsen when the fees are getting higher. However, this is the case for many or most opponents' strategies, and definitely for that of our "main" competitor, Mertacor which was also the 2019 champion.

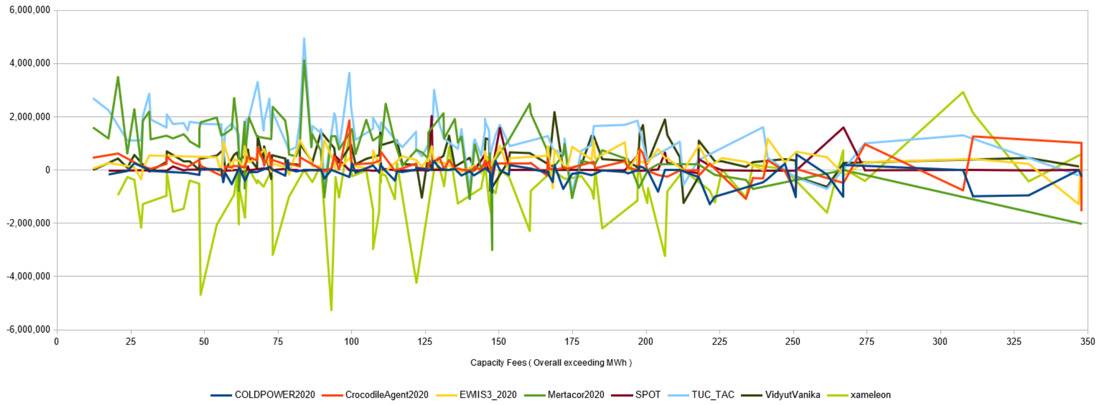

Figure 5.15: Scores of type 5 Regular games

The results for the 5-player games can be seen in Figure 5.15, and are similar to the 8-player ones. TUC-TAC's strategy seems to be the best when the exceeding energy is less than 150 MWh.

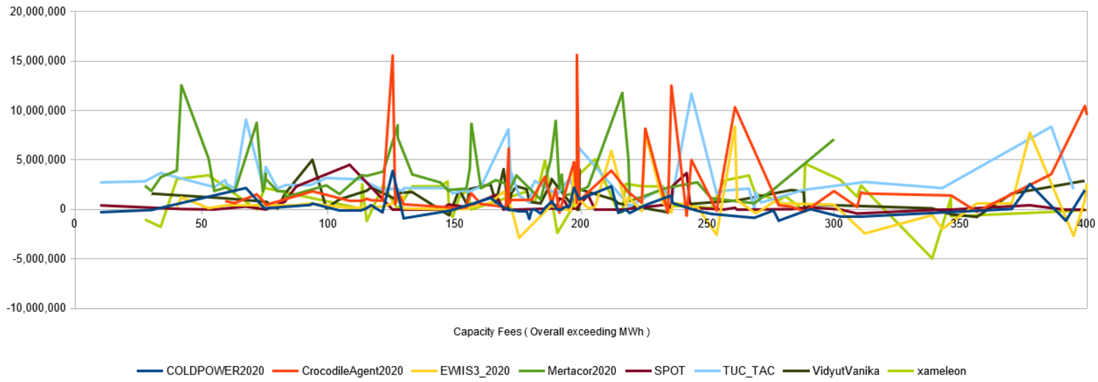

Figure 5.16: Scores of type 3 Regular games

Then, Figure 5.16 cannot provide us with any useful information about the association of the scores with the capacity fees in the case of 3-agent games. The main reason this is happening is the huge variance of the outcomes of different triplets of contestants.

In conclusion, after studying the previous graphs depicting scores in relation to capacity fees, we can most safely infer that TUC-TAC's retail market strategy achieved its goals. As mentioned in the previous chapter, this strategy was created to mitigate the costs of the transmission capacity fees across more than one agent while TUC-TAC could still take the highest share of the tariff profits. This strategy worked exceptionally well when the majority of the agents were part of the game, specifically in 8 and 5 player games. At the same time, this strategy provided TUC-TAC with a very profitable stable performance throughout the 3 player games, but it was not enough to allow it to emerge as the winner in that category.

5.3.4 TUC-TAC's Tariff Profits

In this section, we will demonstrate how each different type of tariff offered by TUC-TAC, assisted to its victory. There are 4 different tariff types offered by TUC-TAC, namely Consumption tariffs, Thermal Storage Consumption Tariffs, Solar Production Tariffs, and Wind Production Tariffs (see section 4.3.1). In figure 5.17 below, we can see the net profits from Consumption and Thermal Storage Consumption tariffs, while in figure 5.18 we can see the losses deriving from the use of Solar Production and Wind Production tariffs.

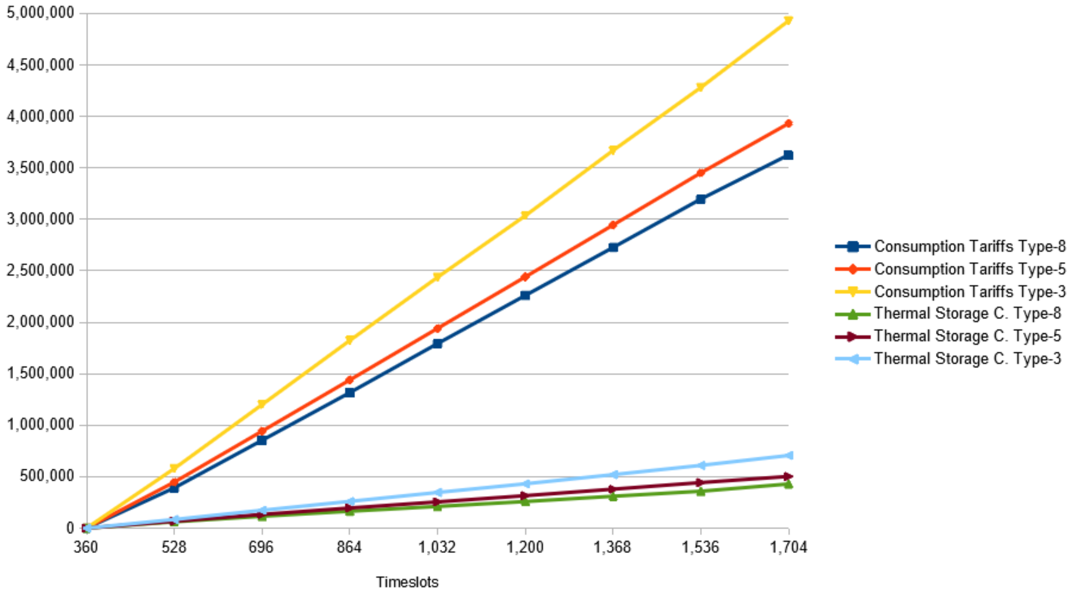

Figure 5.17: Average profits through the course of a game

As observed, the main source of income of TUC-TAC comes from the Consumption tariffs, while a smaller but considerable portion of it is the result of Thermal Storage Consumption tariffs.

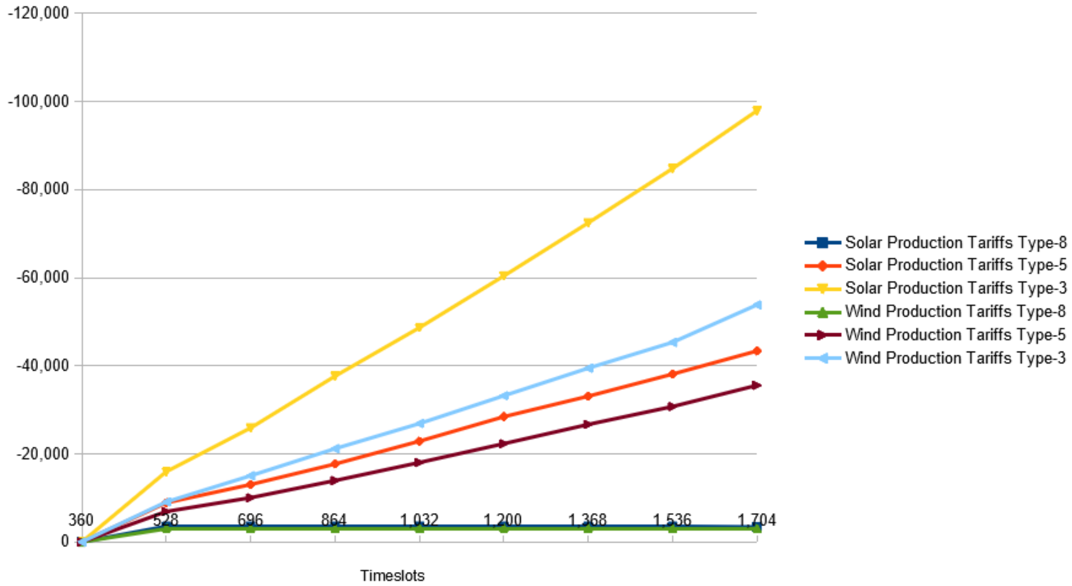

Figure 5.18: Average Losses through the course of a game

Figure 5.18 can not show us the real effect that renewable energy has on PowerTAC. The only thing we can see is the amount of money spent in each case to acquire the useful effects of that power type, besides that, it is visible that the losses in each case are very small to be considered harmful. Some of these useful effects of Production Costumers, have to do with the transmission capacity fees. Specifically, when calculating the fees that each agent has to pay, the Balancing Utility of PowerTAC charges each agent according to its contribution to the net demand. So, if an agent has customers that produced energy in that timeslot, that will greatly reduce the transmission capacity fees that the agent will have to pay. In addition to that, it was necessary to compete and increase the tariff prices for production customers, especially in 3-player games, because other agents like EWIIS3_2020 had increased amounts of profits when they could get low-cost energy. Also, an agent can sell or provide immediately to its customers the produced energy, but this technique usually does not generate enough profit. Of course, it is possible to have sustainable energy tariffs as the main source of income in a game, like some other brokers tried in this year's finals, but to achieve that we would need a whole different approach.

It is clear now how important the Consumption Tariffs are for the viability of a PowerTAC agent. At the same time, there are other sources of income that are not equally important but can be considered when deciding which tariff types to offer. In our case, we found out that the Thermal Storage Consumption tariffs can also be key to making substantial profits in a game (Figure 5.17).

Chapter 6

Conclusions

The importance of the "smart" grid to achieve a coal emission-free society is indisputable, thus the need to test the aspects of the new Smart Grid before deploying is essential. This is why simulation environments such as PowerTAC are important. This Thesis presented the strategy of TUC-TAC 2020 which was the Champion of The PowerTAC 2020 competition. We argued that the novelty and success of the TUC-TAC's strategy lies in its basic principle, which was applied in the retail market for the first time in PowerTAC. That principle resembles to some extent an equilibrium strategy of the Lemonade Stand multiagent zero-sum repeated game, in which the agents try to only acquire the half of the available profit, thus opponents share the other half. However because of the nature of the problem it is difficult, in our view, to solve for an actual equilibrium strategy for PowerTAC.

6.1 Future work

The success of TUC-TAC was very good news for our team which worked a lot to achieve this. However, we believe that our agent has still a lot of room for improvement before the next tournament in 2021. As anyone can easily infer by reading the Experiments and Results part of this thesis, the wholesale market part of TUC-TAC has certain drawbacks that need to be overcome. Our first priority is to implement and add a wholesale market limit price predictor. Along with the predictor, the Monte Carlo Tree search part of the wholesale module needs to be reworked and re-tuned in sight of knowledge accumulated during the 2020 finals.

At the same time, it is very important to not forget to improve the retail market module. It is only reasonable that our opponents will try to "counter" our strategy by finding flaws in our approach. Therefore we will look for new ways to improve our agent in the retail market too, having as a first priority to reduce the Transmission capacity fees as much as possible. The first step to do that will be the addition of a net demand predictor which will be able to identify the upcoming dangerous demand peaks and warn our agent in time. All the necessary preparation for this predictor has already been made, which means an improved version of TUC-TAC is already on sight. Also, we are looking into ways to support more tariff types in a way that would be profitable for our agent. Currently, we are looking into Electric Vehicles and Storage Tariff types with the purpose of adding them to our main tariff strategy. By doing that we expect to have increased profits compared to that of the other competitors.

In the end, we can infer that the PowerTAC is only the beginning since the techniques that were developed for TUC-TAC can also be applied in other multi-agent domains. For instance, the generic "equilibrium" strategy we come up with for the Retail Market of this competition is conceivably a simple but powerful strategy to use in a host of different competitive game domains, as well.

References

[1] W. Ketter, J. Collins, and M. d. Weerdt, "The 2020 power trading agent competition (march 30, 2020).," *ERIM Report Series Reference No. 2020-002,*, 2020. vii, 10, 12, 15, 16

[2] A. Sykulski, A. Chapman, E. Munoz de Cote, and N. Jennings, "The winning strategy for the inaugural lemonade stand game tournament," *Frontiers in Artificial Intelligence and Applications*, vol. 215, 01 2010. vii, 2, 25, 32

[3] C. M. M. Porag, C. Kiekintveld, S. Tran, and W. Yeoho, "Bidding in Periodic Double Auctions Using Heuristics and Dynamic Monte Carlo Tree Search," 2018. 3, 20, 24, 37, 38, 40

[4] S. Massoud Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power and Energy Magazine*, vol. 3, no. 5, pp. 34–41, 2005. 4

[5] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid — the new and improved power grid: A survey," *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 944–980, 2012. 4, 8

[6] C. Akasiadis and G. Chalkiadakis, "Decentralized large-scale electricity consumption shifting by prosumer cooperatives," *ECAI 2016*, p. 157–183, 2016. 5, 7

[7] A. Angelidakis and G. Chalkiadakis, "Factored mdps for optimal prosumer decision-making in continuous state spaces," in *Multi-Agent Systems and Agreement Technologies* (M. Rovatsos, G. Vouros, and V. Julian, eds.), (Cham), pp. 91–107, Springer International Publishing, 2016. 5

[8] A. Angelidakis and G. Chalkiadakis, "Factored mdps for optimal prosumer decision-making in continuous state spaces," EUMAS, 2015. 5

[9] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings, "Putting the 'smarts' into the smart grid: A grand challenge for artificial intelligence," *Commun. ACM*, vol. 55, p. 86–97, Apr. 2012. 6

[10] G. Chalkiadakis, V. Robu, R. Kota, A. Rogers, and N. Jennings, "Cooperatives of distributed energy resources for efficient virtual power plants," in *The Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011) (06/05/11)*, pp. 787–794, May 2011. Event Dates: 2-6 May 2011. 6

[11] R. Kota, G. Chalkiadakis, V. Robu, A. Rogers, and N. R. Jennings, "Cooperatives for demand side management," in *The Seventh Conference on Prestigious Applications of Intelligent Systems (PAIS @ ECAI) (30/08/12)*, pp. 969–974, August 2012. 7

[12] A. Jenn, K. Springel, and A. R. Gopal, "Effectiveness of electric vehicle incentives in the united states," *Energy Policy*, vol. 119, pp. 349 – 356, 2018. 7

[13] A. Koufakis, E. S. Rigas, N. Bassiliades, and S. D. Ramchurn, "Towards an optimal ev charging scheduling scheme with v2g and v2v energy transfer," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 302–307, 2016. 7, 8

[14] E. S. Rigas, S. D. Ramchurn, and N. Bassiliades, "Managing electric vehicles in the smart grid using artificial intelligence: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, pp. 1619–1635, Aug 2015. 8

[15] W. Ketter, J. Collins, and P. Reddy, "Power tac: A competitive economic simulation of the smart grid," *Energy Economics*, vol. 39, pp. 262 – 270, 2013. 9

[16] J. Casazza and F. Delea, *Understanding Electric Power Systems: An Overview of Technology, the Marketplace, and Government Regulation.* 06 2010. 10

[17] J. Serrano Cuevas, A. Y. Rodriguez-Gonzalez, and E. Munoz de Cote, *Fixed-Price Tariff Generation Using Reinforcement Learning*, pp. 121–136. Cham: Springer International Publishing, 2017. 20

[18] G. Susobhan, E. Subramanian, S. P. Bhat, S. P. Gujar, and P. Paruchuri, "VidyutVanika: A Reinforcement Learning Based Broker Agent for a Power Trading Competition," 2019. 20, 41

[19] S. Ozdemir and R. Unland, "The Strategy and Architecture of a Winner Broker in a Renowned Agent-Based Smart Grid Competition," *Web Intelligence*, p. 165–183, 2017. 20, 29, 30, 40

[20] S. Ozdemir and R. Unland, "AgentUDE17: A Genetic Algorithm to Optimize the Parameters of an Electricity Tariff in a Smart Grid Environment," 2017. 20

[21] D. Urieli and P. Stone, "TacTex'13: A Champion Adaptive Power Trading Agent," 2013. 20

[22] F. Natividad, R. Y. Folk, W. Yeoh, and H. Cao, "On the use of off-the-shelf machine learning techniques to predict energy demands of power tac consumers," in *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets* (S. Ceppi, E. David, C. Hajaj, V. Robu, and I. A. Vetsikas, eds.), (Cham), pp. 112–126, Springer International Publishing, 2017. 20

[23] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of monte carlo tree search methods," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, pp. 1–43, March 2012. 23

[24] I. Szita, G. Chaslot, and P. Spronck, "Monte-carlo tree search in settlers of catan," in *Advances in Computer Games* (H. J. van den Herik and P. Spronck, eds.), (Berlin, Heidelberg), pp. 21–32, Springer Berlin Heidelberg, 2010. 24

[25] A. Theodoridis and G. Chalkiadakis, "Monte carlo tree search for the game of diplomacy," in *11th Hellenic Conference on Artificial Intelligence*, SETN 2020, (New York, NY, USA), p. 16–25, Association for Computing Machinery, 2020. 24

[26] P. Ciancarini and G. P. Favini, "Monte carlo tree search in kriegspiel," *Artificial Intelligence*, vol. 174, no. 11, pp. 670 – 684, 2010. 24

[27] V. Robu, R. Kota, G. Chalkiadakis, A. Rogers, and N. Jennings, "Cooperative virtual power plant formation using scoring rules," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, Jul. 2012.

[28] F. Bistaffa, A. Farinelli, G. Chalkiadakis, and S. D. Ramchurn, "Recommending fair payments for large-scale social ridesharing," in *Proceedings of the 9th*

ACM Conference on Recommender Systems, RecSys '15, (New York, NY, USA), p. 139–146, Association for Computing Machinery, 2015.