Sustainable Cities and Society 49 (2019) 101590

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

A competitive and profitable multi-agent autonomous broker for energy markets

Ansel Y. Rodríguez González[a,b,*], Miguel Palacios Alonso[e], Fernando Lezama[c], Lino Rodríguez[e], Enrique Muñoz de Cote[e], Eduardo F. Morales[e], Luis Enrique Sucar[e], Douglas D. Crockett[e,d]

[a] Consejo Nacional de Ciencia y Tecnología, Ciudad de México, México
[b] Unidad de Transferencia Tecnológica, Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE-UT3), Tepic, México
[c] GECAD Research Group, School of Engineering, Polytechnic of Porto, Portugal
[d] Peace Corps México, Avenida Universidad, Santiago de Querétaro, Querétaro,México
[e] Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla, Puebla, México

ARTICLE INFO

Keywords:
Energy markets
Simulation
Autonomous agent

ABSTRACT

Free and competitive energy markets are a recent and increasing phenomenon in several countries. Understanding these new energy markets and estimating their possible evolutions are current challenges of the research community. To avoid real market risks, the research community has developed autonomous traders and tested them in the Power Trading Agent Competition (Power TAC), a sophisticated energy market simulator. In this paper, we present COLDPower'16, a competitive autonomous trader composed of expert agents in specific kinds of markets and customers that combines local strategies into a global strategy to maximize profit. The local strategy of each tariff expert agent uses reinforcement learning algorithms, while the local strategy of the wholesale expert agent estimates future energy prices and the amount of energy that can be negotiated to buy energy when prices are low and sell energy when prices are high. COLDPower'16 was tested in Power TAC 2016. It achieved 2nd place in the final round of this international competition with 7 autonomous agent brokers.

1. Introduction

One step in the way to sustainable cities and society is the massive penetration of electric vehicles and renewable energy (RE) using as primary sources wind and the sun. On the other hand, the use of smart technology with the Internet of Things (IoT) gives consumers new ability to change their behavior aiming at a more efficient management of energy in their homes and offices (Lezama, Palominos, Rodríguez-González, Farinelli, & Munoz de Cote, 2017; Paiho et al., 2018). In this context, free and competitive energy markets are a recent and increasing phenomenon in several countries (Fanelli, Maddalena, & Musti, 2016; Imran & Kockar, 2014; Lezama et al., 2018; Vargas, Saavedra, Samper, Rivera, & Rodriguez, 2016). Understanding these new energy markets and estimating their possible evolution are current challenges of the research community (Golmohamadi, Keypour, Bak-Jensen, & Pillai, 2019; Mirakhorli & Dong, 2018; Zou, Chen, Yu, Xia, & Kang, 2017). The internal variables of energy markets depend on many external variables. Also, the supply and demand of energy, for example, can be influenced by weather conditions, big social events (like Olympic games), or government incentives. As a consequence, the strategies of trading brokers must adapt to the changes.

Due to the high-cost of testing brokers strategies in real energy markets, it is essential to use a reliable simulation environment that can emulate real market conditions based on the laws of supply and demand. To avoid real market risks, the research community has developed autonomous traders (Buljevic et al., 2012; Chowdhury, 2016; Chowdhury, Folk, Fioretto, Kiekintveld, & Yeoh, 2017; Diamantopoulos, Symeonidis, & Chrysopoulos, 2013; Hoogland & La Poutré, 2017; Kuate, Chli, & Wang, 2014; Kuate, He, Chli, & Wang, 2013; Liefers, Hoogland, & La Poutré, 2014; Matetic, Babic, Matijas, Petric, & Podobnik, 2012; Ntagka, Chrysopoulos, &0 Mitkas, 2014; Serrano Cuevas, Rodriguez-Gonzalez, &1 Munoz de Cote, 2017; Urban & Conen, 2017; Urieli & Stone, 2014, 2016) and has tested them in the Power Trading Agent Competition (Power TAC) (Ketter, Collins, & Reddy, 2013), a complex energy market simulator that:

- Accounts for a broad set of the attributes presented in a real energy market.

* Corresponding author at: Consejo Nacional de Ciencia y Tecnología, Ciudad de México, México.
E-mail addresses: ansel@cicese.mx (A.Y. Rodríguez González), mkpalacioc@inaoep.mx (M. Palacios Alonso), flzcl@isep.ipp.pt (F. Lezama), linobi@inaoep.mx (L. Rodríguez), emorales@inaoep.mx (E.F. Morales), esucar@inaoep.mx (L. Enrique Sucar), douglas@inaoep.mx (D.D. Crockett).

https://doi.org/10.1016/j.scs.2019.101590
Received 22 August 2018; Received in revised form 2 May 2019; Accepted 2 May 2019
Available online 21 May 2019

- Considers several types of consumption and production customers with predictable behaviors
- Includes simulated weather condition models.
- Provides a comprehensive reward and penalty framework to promote smart energy trading practices.

Power TAC also stands for an annual tournament (Ketter, Collins, & Weerdt, 2016), where traders designed by research teams from around the world compete against each other.

In this paper, we present COLDPower'16, a competitive autonomous trader composed of expert agents in specific kinds of markets and customers that combines local strategies into a global strategy to maximize profit. COLDPower'16 includes three separate and integrated modules:

1. The **data management module** provides timely and preprocessed information to expert agents.
2. The **retail market module** includes the production and consumption tariff expert agents. The production tariff expert agent tries to attract production customers, while the consumption tariff expert agent tries to attract consumption customers. It is important to note that any imbalance between what is produced and consumed by the customers of an agent generates additional costs and are penalized in the PowerTAC Competition. The local strategy of each tariff expert agent uses reinforcement learning techniques (Watkins & Dayan, 1992). In order to maximize the overall profit, each of these two independent optimization processes perform local improvement steps related to publishing consumption and production tariffs.
3. The main component of the **wholesale market module** is the wholesale market expert agent. Although the wholesale market can be seen simply as an extra component to balance the broker, it also provides a business opportunity for the broker because it can buy energy when prices are low and sell energy when prices are high. The wholesale market expert agent exploits this business opportunity by recognizing the energy imbalance that tariff expert agents may cause. To maximize profits, this agent estimates future energy prices and the amount of energy that can be traded; decides if it is time to buy, sell, or hold; and also decides the amount of energy to trade.

COLDPower'16, our multi-agent autonomous trader, was tested in Power TAC 2016. COLDPower'16 achieved 2nd place in the final round of this international competition with 7 autonomous traders.

The outline of this paper is as follows: In Section 2 energy markets are described. Section 3 provides a brief introduction to Power TAC and 4 describes the broker agents developed for it. In Section 5 our broker agent COLDPower'16 is presented. Section 6 shows the final round results of Power TAC 2016 and compares the COLDPower'16 achievements with the other competing traders. Finally, in Section 7 some conclusions and future work are discussed.

2. Energy markets

A market can be defined as a dynamic environment in which brokers complete transactions by buying and selling assets. The market establishes a mechanism and a fixed set of rules to facilitate competition between brokers (Martinot, Chaurey, Lew, Moreira, & Wamukonya, 2002). An energy market, like any other market, is divided by the type of broker transactions. Low volume and high price transactions define the retail market for example, whereas high volume and low price transactions define the wholesale market. An additional market, called the balancing market, relates to the specific nature of electric energy. We describe these three markets, as modeled by Power TAC, in the next subsections.

2.1. Retail market

Brokers compete in this market by using tariffs to gain customer subscriptions (either consumers or producers). A tariff is a contract between a broker and a producer or a consumer to trade an amount of energy under certain conditions (Wissner, 2011). Such conditions might include the price per energy exchanged, minimum subscription time, early withdraw payments, and others (Mont & Turner, 1999). Customers at the retail market periodically evaluate such tariffs to decide which represent the best subscription option and to replace previous contracts they have with other brokers. Because of this, a broker should explore ways to create new tariffs and modify or revoke existing tariffs in a smart way. Once a broker decides to publish a new tariff, that tariff is broadcast to all customers, and also to other brokers in the retail market. Customers can analyze the published tariff and decide to subscribe. Brokers can also react to such a change by creating news tariffs to attract customer subscriptions.

2.2. Wholesale market

The simulation environment of Power TAC wholesale market is similar to existing wholesale markets in Europe or North America, but is simplified and adapted according to certain rules that do not follow any specific market structure of existing electricity markets (Ketter et al., 2016). In fact, the wholesale market in Power TAC can be seen as an abstraction of day-ahead and intra-day wholesale markets, in which the brokers are allowed to put continuous bids (each hour) for future energy delivery between 1 and 24 h in advance. Due to the maximum horizon of 24 h, the authors of (Ketter et al., 2016) mention that this market might be called a "day-ahead market"[1] Brokers interact with large generating facilities and other participants in the wholesale market, and therefore, the price of energy depends exclusively on the energy supply and demand at the end of each auction. In general, the price of energy at the wholesale market is lower compared to prices in the retail market due to the larger volume of energy traded. Most of the energy handled in the wholesale market is traded in this day-ahead/intra-day scheme, achieving a balance between supply and demand. However, deviations may still take place between the closing of the wholesale market and the delivery (one hour in advance in power TAC market design). To solve these situations, a balancing market is included in which buyers and sellers can trade volumes close to real time.

2.3. Balancing market

The balancing market is a type of market created by uncertainties associated with unpredictable factors like weather in the energy grid. Despite the efforts of brokers to maintain a balance between energy sold and bought, uncertainties hinder a perfect equilibrium between brokers, leading to accumulated energy credits and debits (Casazza & Delea, 2003). Regardless of the brokers' behavior, a balance of energy must be achieved to ensure overall grid stability. This scenario opens the possibility for some entities to provide balancing services to brokers for a fee, creating the balancing market.

3. Power TAC

Power TAC is a complex energy market simulator (Ketter et al., 2013). Fig. 1 shows a block diagram of its architecture. A set of brokers compete against each other in Power TAC by using a multi-agent

[1] It is worth noting that power TAC wholesale market does not follow the rules of most "day-ahead" markets in which the trade horizons are typically 6-30 h in the future and a single daily clearing (see Imran & Kockar, 2014). Instead, power TAC wholesale market design is a simplified abstraction that combines day-ahead and intra-day market characteristics.

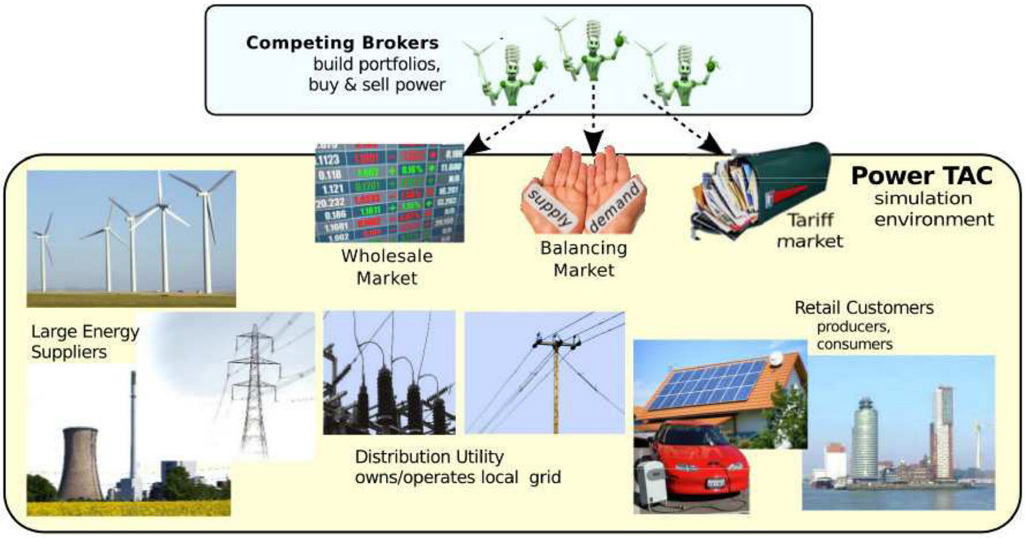

Fig. 1. Major elements of the Power TAC scenario (Ketter et al., 2016).

approach (Maenhoudt & Deconinck, 2010), where brokers can buy or sell energy to their customers in three different markets: the tariff market, the wholesale market and the balancing market. The wholesale market is similar to the Nord Pool market in Scandinavia or the FERC (Federal Energy Regulatory Commission) markets in North America (Ketter et al., 2016), but Power TAC models them all as a single region and includes an abstraction of day-ahead/intra-day market (see Section 2.2). The PowerTAC wholesale market, unlike the Nord Pool Spot market, uses day-ahead/intra-day "continuous" trading mechanism. In such a mechanism, brokers can submit orders to the wholesale market for delivery between one and 24 hours in the future. The time unit used in Power TAC is a "times-lot", which represents one simulated hour. When the simulation clock is advanced to a new time slot, the wholesale market clears the orderbook for each of the enabled time slots.

The brokers can publish tariffs at any given time-slot. After publishing a tariff, the customers can evaluate the offers and decide if they want to stay with the same tariff or change to any other available tariff which may belong or not to the same broker. The objective of each broker is to publish attractive tariffs, so that the producing-customers want to sell energy to it, and the consuming-customers want to buy energy from it. If a broker exceeds or cannot fulfill its energy commitments, either at the wholesale market or at the retail market, it is charged for any imbalance it may have caused, and this charge is deducted from its income. At the end of every time-slot, each broker receives a profit that depends on the incomes, expenses and imbalance fees charged. Since this is a profit-oriented simulation, an international competition has been held every year since 2012, where many teams submit their brokers to be tested against other brokers for research purposes. In order to promote this research, Power TAC keeps detailed logs of each competition.

4. Broker agents

This section provides a look to the agent brokers implemented for trading energy in the PowerTAC simulator. Like PowerTAC has evolved with each edition, the agent brokers implemented by each team have also done so. Not all strategies followed by the agents and their versions have been published, but most of them are available (e.g. CrocodileAgent, Mertacor, SPOT, AgentUDE and Maxon).

On the retail market, CrocodileAgent (Buljevic et al., 2012; Matetic et al., 2012) replaces tariffs that are not profitable enough with tariffs that have higher utility. Tariffs that show insufficient utility over time are revoked completely. CrocodileAgent offers time of use tariffs with different expiration times but with specific, fast-profit generating parameters (i.e., periodic payment). The cleared prices from the wholesale market and the amount of energy traded by these prices are used to get the minimum price of energy for tariff specification rates which will not cause and losses if applied to a certain rate. CrocodileAgent predicts the amount of energy customers will need in the next 24 timeslots and estimates clearing prices on the wholesale market by means of a triple exponential smoothing technique better known as the Holt-Winters Method. CrocodileAgent computes the price offered on the wholesale market depending on the number of tries left. As the respective timeslot approaches, the price that it is willing to pay in order to buy the needed amount of energy is increased. If there is only one try left, the broker offers a predefined maximum price.

Mertacor (Diamantopoulos et al., 2013; Ntagka et al., 2014) implements two types of strategies for the retail market: a tariff formation strategy and a tariff update strategy. Both strategies are treated as optimization problems, where the objective function maps to the broker's maximum profit, while retaining an adequate portion of the customers market share. The optimization problems are solved using Particle Swarm Optimization (PSO) techniques. For the wholesale market the following policies are designed: an adaptive price formation policy, a policy for forecasting energy consumption that employs Time Series Analysis primitives, and two shout update policies, a rule-based policy that acts rather hastily, and one based on Fuzzy Logic that describe the agent's state in terms of the number of consumers (producers) under contract and the portfolio balance.

SPOT (Chowdhury, 2016; Chowdhury et al., 2017) uses an unsupervised reinforced learning technique in order to publish the optimal tariff to gain both the most subscribers and the greatest net balance. To achieve this goal the problem is modeled as a Markov Decision Process (MDP) and the Q-Learning algorithm is used to discover the optimal policy. For the wholesale market SPOT uses decision trees and neural networks to predict the clearing price. Other features like weather forecast and time of day are also included for training the predictors.

AgentUDE (Ozdemir & Unland, 2015) offers two kind of consumer tariffs in the retail market. If the offered price is the cheapest one among other tariffs, then it is published with a early withdrawal penalty fee. Otherwise, early withdrawal penalty fee is not set and tariff value is adjusted considering the cleared prices of the wholesale market and distribution fee. On the wholesale market a price prediction takes place in two steps: The base price is a predicted utilizing past data. Afterwards, the final price is differentiated using the base price. Recently, a new AgentUDE version (Özdemir & Unland, 2018) was presented which use an online genetic algorithm that optimizes the parameters of an electricity consumption tariff, such as unit retail price, periodic, sign-up, and early withdrawal penalty payments on the fly. Additionally, it uses time-of-use (TOU) price scheme to reduce peak- demand charges.

Maxon (Urban & Conen, 2017) uses four kinds of tariffs in the retail market: flat, tiered, TOU and tariffs for controllable devices. Flat tariffs are offered to customer models who cannot, or do not want to, alter their behavior. This kinds of tariff is based on the baseline price which is computed by using the costs the broker had to pay for each kWh during the last two weeks. Tiered and TOU tariffs allow to flatten peak demands by rewarding customers if they alter their usage behavior. To generate the TOU tariffs Maxon implements a method, which scale the prices of the maximum and minimum demands, that uses the baseline price, two scaling factors and an inner Hill Climbing Algorithm to get the minimum and maximum over the aggregated bootstrap usage data of all customer models. The main purpose of tariffs for controllable devices is to offer balancing capacities to the Distribution Unit. On the wholesale market, Maxon buys the needed energy as early as possible. A multiple linear regression model it is used to predict the amount of needed energy. Maxon splits the order into multiple smaller orders at different price levels. If it does not get the needed energy the price limit is increased, based on the last clearing prices, up to a level until it is cheaper to run into balance.

CWIBroker (Hoogland & La Poutré, 2017; Liefers et al., 2014) try to balancing its demand and supply in each timeslot by estimating its customers' prosumption. To do that it uses a linear regression model, which takes into account historical prosumption data and weather data. On the wholesale market, CWIBroker bids its entire estimated demand in the first auction round. In the remaining auction rounds it bids the estimated demand it has not yet acquired in the previous auction rounds. It use a heuristic to compute its limit price that for each auction round, the limit price for the next timeslot is computed from the worst case limit price in the most recent timeslot. On the retail market, CWIBroker publishes tariffs with an energy unit price lower than the competition price, provided that the unit price is higher than an estimation of the cost price. The strategic consist in mostly decreases its tariff price, which depends on the consumption of the broker's customers divided by the consumption of all consumers in the simulation.

Other agent brokers implemented for trading energy in the PowerTAC simulator are AstonTAC and TacTex. AstonTAC (Kuate et al., 2013, 2014) uses a Non-Homogeneous Hidden Markov Model to forecast energy demand and price while TacTex (Urieli & Stone, 2014, 2016), uses a modified version of Tesauro's bidding algorithm, where they modeled the sequential bidding process as a Markov Decision Process for the wholesale market.

5. COLDPower 2016: a multi-agent autonomous broker

COLDPower'16, is a competitive autonomous trader composed of expert agents in specific kinds of markets and customers that contribute with local strategies to a global strategy that maximizes profit. It includes three separate and integrated modules: data management, retail market and wholesale market. Fig. 2 shows the COLDPower'16 architecture.

The data management module provides an abstraction of the server communication to the other agents. It also provides timely information to the market experts. Server messages are compiled in the data storage. The experts do not have direct access to the data storage, but only through predefined views. The views include simple structures like private data, statistics, ordered sets, and others that can be easily used by the experts.

The retail market module includes the software components for the production and consumption tariff experts. The production tariff expert tries to attract production customers, while the consumption tariff expert tries to attract consumption customers. It is important to note that the imbalance between what is produced and consumed by customers of an agent generates additional costs, which are penalties in the PowerTAC competition. To attract customers, the experts define contracts, which are known as tariffs in the retail market. The main tariff variables include the price per kilowatt (kw), subscription bonuses, forced deadlines, and penalties for anticipated unsubscription. Tariffs generated are sent to the server for publication. Customers may assess the published tariffs at any time and switch to another tariff or stay in it. The strategies of both tariff experts are detailed in Section 5.1.

The main component of the wholesale market module is the wholesale market expert. In this market, also known as the day ahead market, agents can negotiate energy for the next 24 hours. In other words, for a each specified target hour in the future, agents in the wholesale market can negotiate energy in any of the previous 24 hours to it. This allows agents to have 24 chances to get extra power to meet the needs of its consumption customers or get rid of excess energy produced by its production customers to avoid additional penalties. Although the wholesale market can be seen simply as an extra component to balance the broker, it also provides a business opportunity for the broker because it can buy energy when the price is low and sell energy when the price is high. The wholesale market expert exploits this business opportunity by recognizing the energy imbalance that tariffs experts may cause. The strategy of the wholesale market expert is detailed in Section 5.2

For a each specified target hour in the future, agents in the wholesale market can trade energy in any of the previous 24 hours to it.

5.1. Retail market strategy

As a market participant, each broker develops a portfolio of customers (producers and/or consumers) and interacts with these customers by offering tariff contracts for buying or selling energy. The broker's goal is to increase the number of subscribed customers by satisfying their energy needs and, simultaneously, to increase its profit. To reach this goal a broker must set tariff parameters such as price, sign up bonus or fee, periodic payment, or early withdraw payments. An optimal setting of these parameters results in a very competitive tariff.

COLDPower's retail (or tariff) market strategy differs from most of the other PowerTAC agent strategies in the use of reinforcement learning. In this sense, COLDPower is similar to SPOT. But unlike SPOT, COLDPower does not use a single MDP but two independent processes. A process tries to attract as many producers as possible. The other process attempts to attract only the necessary and sufficient consumers to reduce the energy imbalance of the client portfolio.

The strategy is based on off-line and on-line reinforcement learning (RL). The pair of MDP models are built to adapt to market tariff changes and used for publishing tariffs. The solution of these models determines two key parameters of one tariff: price and sign up payment. To maximize the overall COLDPower'16 profit, each of these two independent optimization processes perform local improvement steps related to publishing consumption and production tariffs.

The overall broker's utility or profit can be defined as follow:

$$U_{B_i} = \sum_{t \in T} (u_t^{B_i}) \tag{1}$$

where U_{B_i} is the broker's utility or profit at the end of the game, $u_t^{B_i}$ is the broker's utility at time t. COLDPower'16 aims to maximize this utility at every decision step using two MDP models.

The modeling approach proposed explicitly considers the behavior of the other market participants. It focuses on generic production and consumption tariffs with fixed prices. The next section describes the market representation on both models and the related actions considered.

5.1.1. Consumption MDP model

The consumption MDP model bases its market representation on two variables: Porfolio Status (PS) and Consumption Price Status (CPS) (Reddy & Veloso, 2011; Serrano Cuevas et al., 2017). The set of actions

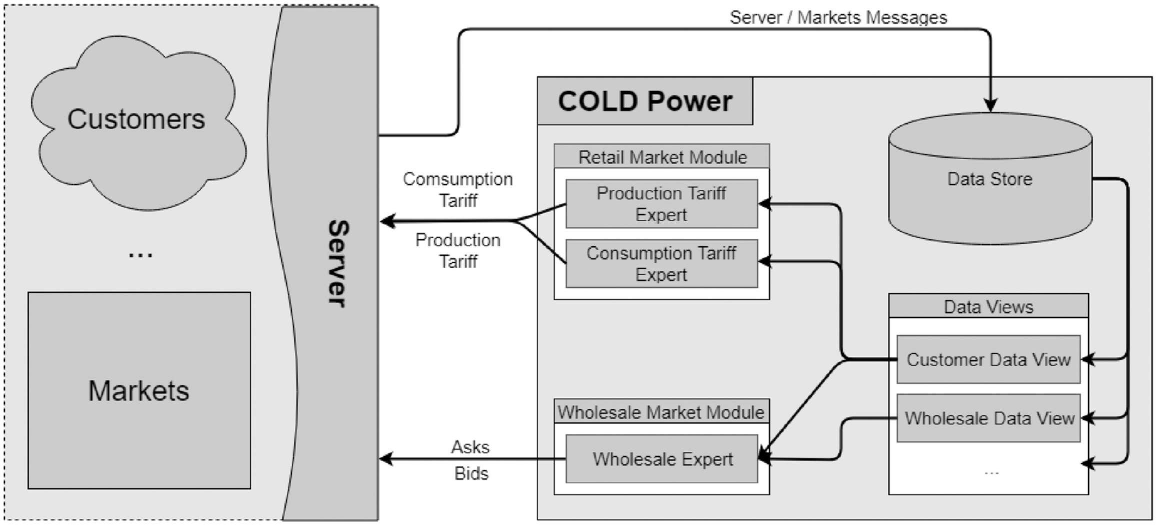

Fig. 2. COLDPower'16 architecture.

is:

$$A = \{\text{up, down, maintain, maxUp, maxDown}\} \quad (2)$$

and the learned policy determines both an optimal price and sign up payment for publishing consumption tariffs.

- State representation.
  The variables that define a state for consumption purposes are:
  - PS = {*shortsupply*, *balanced*, *oversupply*} represents the status of the COLDPower's portfolio in terms of energy imbalance.
  - CPS = {*highOut*, *mediumOut*, *lowOut*, *near*, *far*, *veryFar*} represents the relationship between COLDPower's consumption price (*price*) and set of the tariff prices defined by competitors.
    Values from *highOut* to *lowOut* indicate that *price* is less attractive than the minimum consumption tariff price of competitors. *HighOut* value indicates that *price* is almost as unattractive as the maximum consumption tariff price of competitors, while *lowOut* value indicates the *price* is just a little less attractive than the minimum consumption tariff price of competitors. On the other hand, values from *near* to *veryFar* indicate that *price* is more attractive than the minimum consumption tariff price of competitors. *Near* value indicates that the *price* is just a little more attractive than the minimum consumption tariff price of competitors while *veryFar* value indicates that attractiveness of *price* is the highest (close to the minimum production tariff price of competitors)
- Actions:
  - *up*. To increase price and bonus subscription so that the resulting tariff is less attractive.
  - *down*. To decrease price and bonus subscription so that the resulting tariff is more attractive.
  - *maintain*. To maintain without changes to the current tariffs so that there is no tariff publication.
  - *maxUp*. To greatly increase the price and bonus subscription which is more aggressive than the *up* action.
  - *maxDown*. To greatly decrease the price and bonus subscription which is more aggressive than the *down* action.
- Reward.
  The reward in the consumption MDP model is defined as the following:

$$U_t^{B_i} = P_{t,C}^{B_i}\psi_{t,C} - P_{t,P}^{B_i}\psi_{t,P} - \theta_t \quad (3)$$

where $P_{t,C}^{B_i}$ is the consumption price in t, $\psi_{t,C}$ is the amount of energy sold, $P_{t,P}^{B_i}$ is the production price, $\psi_{t,P}$ the amount of produced energy, and θ_t is the cost generated by imbalance. In the PowerTAC context this utility measure is the cash amount (in t) in its bank account.

5.1.2. Production MDP model

The MPD's solution determines an optimal price for publishing production tariffs.

- State representation.
  The state of the production MDP model is defined by these variables:
  - PPS = {*out*, *highNear*, *mediumNear*, *lowNear*, *far*, *veryFar*} represents the relationship between the COLDPower's production prices and a reference price defined by prices of competitors in an analogous way that described for the consumption model.
  - CR = {*low*, *medium*, *high*, *all*} represents the discretized percentage of production customers subscribed.
  - PS = {*shortsupply*, *balanced*, *oversupply*} represents the status of the COLDPower's portfolio in terms of energy imbalance.
- Actions:
  - *up*. To increase both price and bonus subscription so that the tariff is more attractive.
  - *down*. To decrease both price and bonus subscription so that the tariff is less attractive.
  - *maintain*. To maintain without changes to the current tariffs so that there is no tariff publication.
  - *maxUp*. To greatly increase the price and bonus subscription which is more aggressive than the *up* action.
  - *maxDown*. To greatly decrease the price and bonus subscription which is more aggressive than the *down* action.
- Reward. The production MDP model uses the energy imbalance as a reward that is defined as the following:

$$U_t^{B_i} = -|\psi_{t,C} - \psi_{t,P}| \quad (4)$$

Negative rewards are assigned to energy imbalance (either oversupply or short supply), so that maximizing the reward results in minimizing imbalance.

5.1.3. Learning models

Both MDP models are learned/updated using the Q-learning updating rule (Watkins & Dayan, 1992) defined as:

$$\hat{Q}(s, a) = (1 - \alpha_t)\hat{Q}_{t-1}(s, a) + \alpha_t[r_t + \gamma\hat{Q}_{t-1}^{\max a}(s, a)] \quad (5)$$

It should be noted that, with discrete states and actions, Q-learning

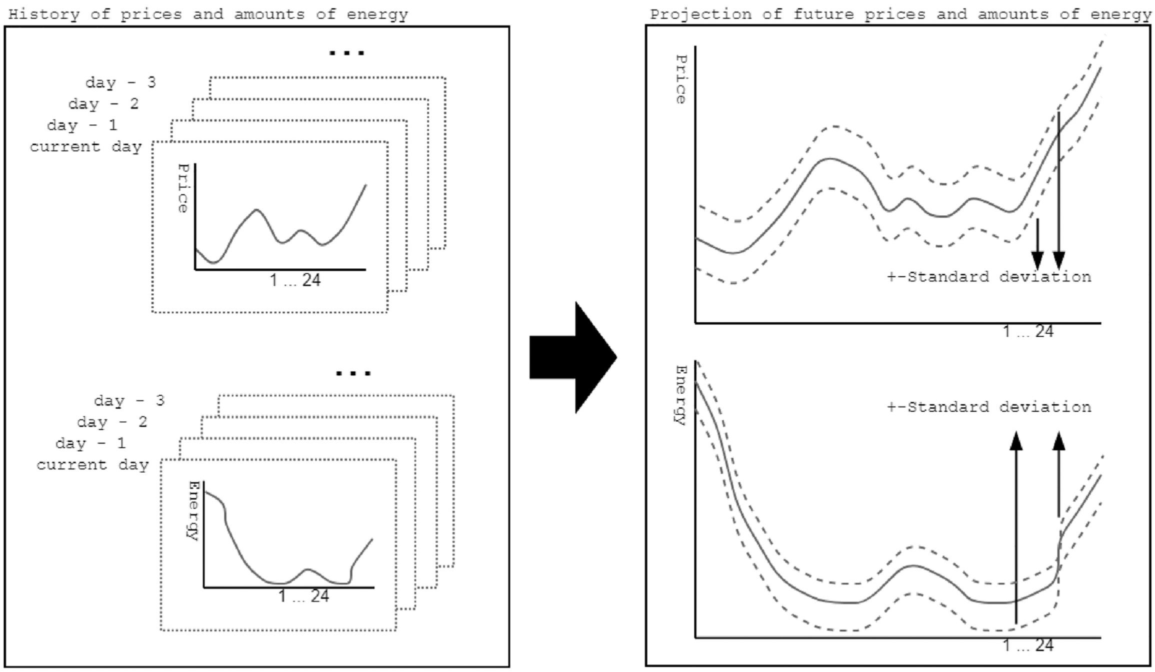

Fig. 3. Estimation of the future energy prices and the amount of energy that can be traded for each 24 opportunities associated with a target time.

converges to the optimum policy with probability one, as long as all the actions are repeatedly sampled in all states (Watkins & Dayan, 1992).

In order to successfully represent the market state, COLDPower learns/updates one Q table for each PowerTAC setting game in terms of number of competitors. In the competition there are three settings with 7, 5 and 3 competitors, so three pairs of models (consumption-production) are learned.

The exploration strategy used is epsilon greedy.

5.2. Wholesale market strategy

Agents in the wholesale market can trade energy for a specified (target) timeslot in any of the previous 24 hours through buy or sale orders. These orders may or may not be executed, depending on the clearing price of the market which is calculated from the limit prices of all orders received by the market (Ketter et al., 2016).

Our proposed wholesale market strategy focuses on exploiting the business opportunity that represents buying energy when the price is expected to be low and selling when the price is expected to be high. It also aims to obtain additional energy needed by consumption customers or to get rid of excess energy provided by production customers in order to balance the broker and to avoid additional penalties.

Like other PowerTAC wholesale agent strategies, this approach allows to diminish the imbalance of the portfolio of clients of the market of tariffs. But it differs from other strategies in that it can work and be profitable by itself even without the existence of a customer portfolio or a strategy in the retail market. In the worst case, if there is a hard competition in the retail market and purchase and sale prices almost intersect, then the best option is not to make the rates more attractive and therefore customers are lost. In this case, the our wholesale market strategy alone can make the broker profitable.

To achieve these objectives the strategy follows these steps for each target timeslot:

1. Obtain an estimated projection of future energy prices and the amount of energy that can be traded for each of the 24 opportunities associated with the target time. The projection is obtained from the history of energy prices and the amounts of energy traded in the market for the target time.
2. Decide, from the estimated projection of the energy price and for each one of the opportunities associated with the target time, whether the best opportunity is to buy, sell, or hold.
3. Compute the price of the buy or sale order for the target time from the estimated projection of the energy price, but only if the current time is an opportunity to buy or sell.
4. Compute the amount of energy to trade in the current time, based on:
   (a) the estimated projection of energy trading for further opportunities,
   (b) the energy of the orders executed for previous opportunities, and
   (c) the estimated energy balance of the customers for the target timeslot.
5. Send an order to the market with the price and energy amount calculated if the current hour is an opportunity to buy or sell.

Steps 1 to 4 are explained below.

5.2.1. Obtaining the estimated future energy prices and amounts of energy that can be traded

Since energy can be traded for any target time in the previous 24 h, and since the hours are repeated every day, we have several repetitions of sets of 24 historical clearing prices and amounts of energy traded for each target time (one per day elapsed of market operation).

Let $P_{h,d} = (p_1, p_2, \ldots, p_{24})$ be a series of 24 historical cleared prices data associated with the target time h of the day d, and $E_{h,d} = (e_1, e_2, \ldots, e_{24})$ be a series of 24 historical amounts of energy negotiated associated with the target time h of the day d. For convenience, we use $p_{h,d,i}$ to denote the p_i cleared price into $P_{h,d}$ and $e_{h,d,i}$ to denote the e_i amount of energy negotiated into $E_{h,d}$.

From the historical cleared prices $P_{h,d^0}, \ldots, P_{h,d^*-2}, P_{h,d^*-1}$, where d^* is the most recent day with a current time greater than the target time, the estimation of the future energy prices $\tilde{P}_h = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{24})$ (Fig. 3) is computed by weighing historical prices. Oldest prices have the smallest weight while recent prices have the highest weight.

For each $i \in \{1, \ldots, 24\}$, therefore:

- $expectedImbalance + negotiatedEnergy > 0$
  - $expectedImbalance + negotiatedEnergy + expectedBuy > expectedSell$
    - $a_k = buy$
      - $expectedImbalance + negotiatedEnergy > expectedSell$
        $e^* = 0$
      - $expectedImbalance + negotiatedEnergy \leq expectedSell$
        $e^* = expectedSell - (expectedImbalance + negotiatedEnergy)$
    - $a_k = sell$
      $e^* = expectedImbalance + negotiatedEnergy + expectedBuy$
  - $expectedImbalance + negotiatedEnergy + expectedBuy \leq expectedSell$
    - $a_k = buy$
      $e^* = expectedSell - (expectedImbalance + negotiatedEnergy)$
    - $a_k = sell$
      $e^* = expectedImbalance + negotiatedEnergy + expectedBuy$
- $expectedImbalance + negotiatedEnergy \leq 0$
  - $-(expectedImbalance + negotiatedEnergy) + expectedSell > expectedBuy$
    - $a_k = buy$
      $e^* - (expectedImbalance + negotiatedEnergy) + expectedSell$
    - $a_k = sell$
      - $-(expectedImbalance + negotiatedEnergy) > expectedBuy$
        $e^* = 0$
      - $-(expectedImbalance + negotiatedEnergy) \leq expectedBuy$
        $e^* = expectedBuy + expectedImbalance + negotiatedEnergy$
  - $-(expectedImbalance + negotiatedEnergy) + expectedSell \leq expectedBuy$
    - $a_k = buy$
      $e^* - (expectedImbalance + negotiatedEnergy) + expectedSell$
    - $a_k = sell$
      $e^* = expectedBuy + expectedImbalance + negotiatedEnergy$

Fig. 4. Cases for computing the amount of energy to trade.

$$\tilde{p}_i = \frac{\sum_{j \in \{0..d^*\}} \alpha^{-j} p_{h,d^*-j,i}}{\sum_{j \in \{0..d^*\}} \alpha^{-j}} \quad (6)$$

where $\alpha \in \{0, \ldots, 1\}$ is a weighting factor. If $\alpha = 1$, then all data, regardless of age, have the same relevance, and $\tilde{p}_i$ is the average of historical cleared prices. When alpha decreases, the relevance of the historical data decreases exponentially regarding its age, in COLDPower, we set $\alpha = 0.9$.

A weighted standard deviation $StdP_h = (stdp_1, stdp_2, \ldots, stdp_{24})$ for $\tilde{P}_h$ is also computed.

$$\mathrm{stdp}_i = \sqrt{\frac{\sum_{j \in \{0..d^*\}} [\alpha^{-j}(p_{h,d^*-j,i} - \tilde{p}_i)]^2}{\sum_{j \in \{0..d^*\}} \alpha^{-j}}} \quad (7)$$

Similarly, the estimated future amounts of energy that can be traded $\tilde{E}_h = (\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_{24})$ and a weighted standard deviation $StdE_h = (stde_1, stde_2, \ldots, stde_{24})$ for $\tilde{E}_h$ are computed (Fig. 3).

5.2.2. Deciding whether to buy, sell, or hold

From the estimated future energy prices $\tilde{P}_h = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{24})$ and for each opportunity to trade i, we decide if it is an opportunity to buy, sell, or hold. To make the smartest decision, the average $\tilde{\tilde{p}}$ and the standard deviation $\mathrm{std}\tilde{p}$ of $\tilde{P}_h$ is computed first. Then a vector $A = \{a_1, a_2, \ldots, a_{24}\}$ of actions with $a_i \in sell, buy, hold$ is built. For all $i \in \{1, \ldots, 24\}$,

$$a_i = \begin{cases} \text{sell} & \text{if } \tilde{p}_i > \tilde{\tilde{p}} + \beta^*\mathrm{std}\tilde{p} \\ \text{buy} & \text{if } \tilde{p}_i < \tilde{\tilde{p}} - \beta^*\mathrm{std}\tilde{p} \\ \text{hold} & \text{othercase} \end{cases} \quad (8)$$

where β, $\beta \geq 0$ is a factor of the standard deviation $\mathrm{std}\tilde{p}$, in COLDPower, we set $\beta = 1$.

5.2.3. Computing the energy price of the buy and sale orders

This calculation is made only for the current timeslot when there is a trading opportunity to *buy* or *sell*.

Let h^* be the current time, the energy price p^* is computed taking, into account the estimated future energy price $\tilde{p}_k$ in $\tilde{P}_h = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{24})$ and the standard deviation $stdp_k$ in $StdP_h = (stdp_1, stdp_2, \ldots, stdp_{24})$ for the time h^*, using the following equation:

$$p^* = \begin{cases} \tilde{p}_k - \gamma^*\mathrm{stdp}_k & \text{if } a_k = \text{sell} \\ \tilde{p}_k + \gamma^*\mathrm{stdp}_k & \text{if } a_k = \text{buy} \end{cases} \quad (9)$$

where γ, $\gamma \geq 0$ is a factor of the standard deviation $stdp_k$, in COLDPower, we set $\gamma = 0.5$.

5.2.4. Computing the amount of energy to trade

This is a critical step for the broker's global strategy because the amount of energy traded is intended to balance the broker. If an excess of energy produced compared to energy consumed is predicted, for example, then wholesale market trades should generate an opposite imbalance in that market to minimize the global imbalance.

First, the expected imbalance (*expectedImbalance*) is computed from the trader's customers portfolio for the target time. Although different prediction techniques can be used, we simply use the value corresponding to the target time in the imbalance profile of the client portfolio. The imbalance profile of the client portfolio is obtained similarly to how the estimated future energy prices are computed. The *expected Imbalance* is positive if the energy produced is greater than the energy consumed, or negative in the opposite case.

Secondly, negotiated energy (*negotiatedEnergy*) is computed from the sum of energy that was bought and sold before the current time for the target time. The negotiated energy is positive if the energy bought is greater than the energy sold, or negative in the opposite case.

Table 1
Parameters used in Power TAC 2016 tournament games (from Ketter et al., 2016).

Parameter	Standard game setting
Length of pregame bootstrap period	14 days
Nominal length of game	60 days
Probability of game end after time slot 1320	$\frac{1}{121}$
Minimum game length	1320
Expected game length	1440
Timeslot length	60 minutes
Time compression ratio	720 (5 s/time slot)
Open time slots on whole sale market	24
Market closing time	1 time slot ahead
Minimum order quantity	0.1 kWh
Meter charge, small customer	[0.01–0.02] €/timeslot
Meter charge, large customer	[0.03–0.10] €/timeslot
Demand assessment interval	168 hours
Peak demand threshold coefficient	[1.0–2.0]
Peak demand charge	[100–1000] €/point
Balancing cost	[0.02–0.06] €/kWh
Slope of regulating market price	10^{-6} €/kWh
Default brokerś min and max bid order prices	−100, −5
Default brokerś min and max ask order prices	0.1, 30
Tariff publication fee	[1000–5000] €
Tariff revocation fee	[100–500] €
Tariff publication interval	6 time slots
Daily bank debt interest rate	4.0%/365...12.0%/365
Daily bank deposit interest rate	0.5 daily bank debt interest rate
Weather report interval	1 hour
Weather forecast interval	1 hour
Weather forecast interval	24 hours

Thirdly, the estimated amount of energy that can be purchased (*expected Buy*), and the estimated amount of energy that can be sold (*expectedSell*) in the future trading opportunities, are calculated from the estimated future amounts of energy that can be traded $\tilde{E}_h = (\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_{24})$.

Finally, the amount of energy to trade (e^*) is computed for the corresponding case (Fig. 4). When the computed amount of energy to trade e^* is 0, no order is sent to either buy or sell.

6. Autonomous brokers competition

COLDPower'16, our multi-agent autonomous trader, was tested in the Power TAC 2016 tournament. Table 1 shows the parameters used by the simulator for standard game settings (Ketter et al., 2016).

As a competitive simulation environment, Power TAC challenges brokers to maximize their profits by buying and selling energy. Brokers compete with each other to attract costumers to maximize their profit at the same time that minimize the costs. Thus, the number of brokers varies during the competition to determine the broker with the best performance during the competition.

The tournament consists of 4 rounds: a qualification round, two seeding rounds and the final round. Only the final round is considered to define the winner of the competition. At the final round, 7 international research institutions participated with their own autonomous brokers. Table 2 shows the name of each broker, institution, and country where it was developed.

Table 2
PowerTAC 2015 tournament competitors.

Broker	Institution	Country
maxon16	Westfaelische Hochschule	Germany
COLDPower	INAOE	Mexico
AgentUDE	Universitaet Duisburg-Essen	Germany
SPOT	UTEP/NMSU	U.S.A.
Mertacor	Aristotle University of Thessaloniki	Greece
AgentCU	The Chinese University of Hong Kong	China
CrocodileAgent	University of Zagreb	Croatia

The final round consisted of 3 blocks of games. The first block included 30 games of 7 brokers, the second block included 63 games of 5 brokers and the third block included 105 games of 3 brokers. Thus, each combination of registered brokers of each size of the game are considered for the same number of games. Regardless of the round, each game's length did not have a fixed number of timeslots, but an expected value of approximately 1440 timeslots, corresponding to 55 days.

For each setting of game in terms of competitors (7, 5 and 3) a pair production and consumption MDP models were trained. The Q tables were initialized with values at 0. Online data from the qualification and seeding rounds were used to train the models. Models learned were used during the final round of the tournament. Only production and consumption tariffs are published.

To measure the performance of each broker, Power TAC shows the results of each block independently by using a standardized broker performance metric as defined in Eq. (10). This metric allows us to compare the results obtained by each broker at each block, even if the number of brokers participating per game changed within each block.

The standardized performance *sp* for broker *b* at block *c* is defined as:

$$\mathrm{sp}_{b,c} = \frac{\mathrm{ap}_{b,c} - \mathrm{avg}_c}{\mathrm{std}_c} \tag{10}$$

where $ap_{b,c}$ is the accumulated profit of broker b at block c, avg_c is the average accumulated profit of all brokers at block c and std_c is the standard deviation of all the brokers at block c.

Table 3 and Fig. 5 show the tournament results in terms of accumulated profits. Table 3 also shows the standardized performance and rank order with COLDPower'16 highlighted.

From Table 3, it can be seen that our agent COLDPower'16 obtained gains in all three scenarios and placed second in each case.

In the 7 and 5 agent scenarios (blocks 1 and 2), which are very competitive scenarios, the COLDPower'16 gains were much higher than most of the other competitors. COLDPower'16 was also close to maxon16, the first place broker. Comparing maxon16 and COLDPower'16 profits in 7 agent scenarios, it can be seen from Fig. 6 that the profit average of COLDPower'16 was greater than the profit

Table 3
PowerTAC 2016 tournament results.

Broker	Accumulated Profits			
	block 1 *size* = 7	block 2 *size* = 5	block 3 *size* = 3	Total
maxon16	28150889	58222359	162300823	248674071
COLDPower	18919929	35255197	86548360	140723486
AgentUDE	9545250	13382258	79616833	102544341
SPOT	-91040	9152999	52695182	61757141
Mertacor	628193	-139322	-404205	84666
AgentCU	-22164073	-22755943	81582603	36662586
CrocodileAgent	-40352748	-78593294	-186380292	-305326333
Broker	**Standarized Performance**			
	block 1 *size* = 7	block 2 *size* = 5	block 3 *size* = 3	Total
maxon16	1.320	1.361	0.611	3.292
COLDPower	0.899	0.796	-0.047	1.648
AgentUDE	0.471	0.258	-0.107	0.621
SPOT	0.031	0.154	-0.341	-0.157
Mertacor	0.064	-0.075	-0.802	-0.814
AgentCU	-0.977	-0.631	-0.090	-1.699
CrocodileAgent	-1.808	-2.005	-2.418	-6.230
Broker	**Rank order**			
	block 1 *size* = 7	block 2 *size* = 5	block 3 *size* = 3	Total
maxon16	**1**	**1**	**1**	**1**
COLDPower	**2**	**2**	**2**	**2**
AgentUDE	**3**	**3**	4	**3**
SPOT	5	4	4	4
Mertacor	4	5	6	5
AgentCU	6	6	**3**	6
CrocodileAgent	7	7	7	7

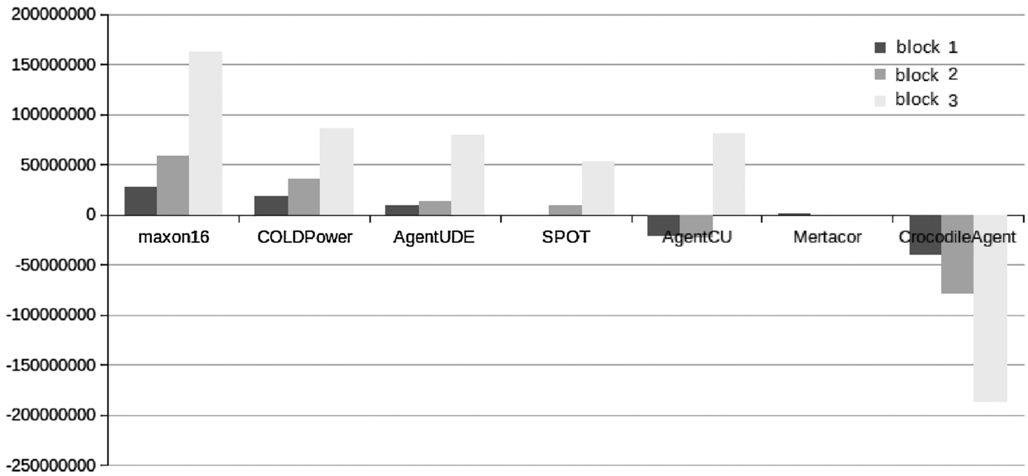

Fig. 5. Accumulated profits by broker for each block.

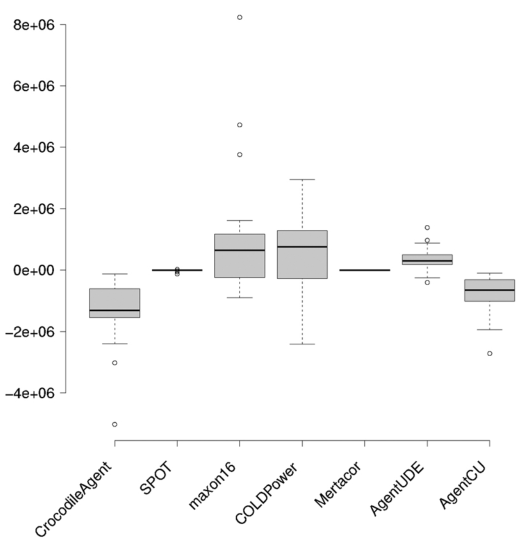

Fig. 6. Accumulated profits by broker for scenario of 7 agents.

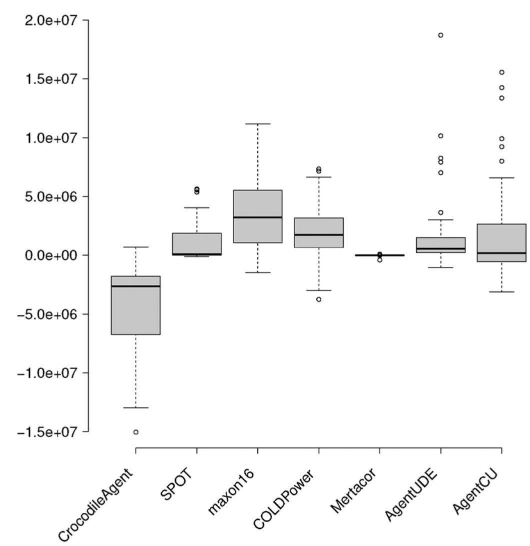

Fig. 8. Accumulated profits by broker for scenario of 3 agents.

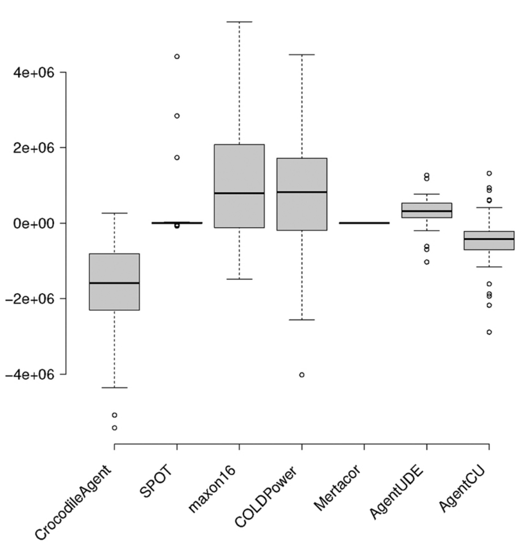

Fig. 7. Accumulated profits by broker for scenario of 5 agents.

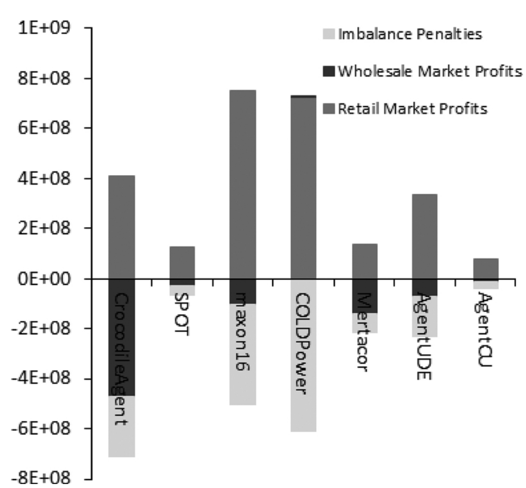

Fig. 9. Main kinds of profits by broker.

average without outliers of Maxon16, while in 5 agent scenarios (Fig. 7) the profit average of COLDPower'16 and maxon16 were very similar.

In the 3 agents scenarios (block 3) COLDPower'16 average gains were less than maxon16 average gains but closer to maxon16 than the other brokers. This does not mean that the earnings of COLDPower16 were scarce, however, since our agent earned over €86 million EUR (Fig. 8).

It is important to highlight that maxon16 not only published

Table 4
Main kinds of profits by broker.

Broker	Tariff market profits	Wholesale market profits	Imbalance penalties
maxon16	749735588	−103502467	−398379484
COLDPower	724176487	5577749	−608869925
AgentUDE	332305551	−73719158	−156289226
SPOT	127888421	−28921373	−37974614
Mertacor	139335817	−140081822	−77835943
AgentCU	76986004	−12209523	−28019660
CrocodileAgent	408478237	−472242219	−239840592

production and consumption rates in the retail market and participated in the wholesale market, but maxon16 also created contracts for storage units and controllable customers (both producers and consumers) in its global strategy, and especially used these costumers in the balance market. Our COLDPower'16 broker did not exploit the opportunity represented by storage units and controllable customers for the balance market which is an extra mechanism to minimize the imbalance penalties.

The retail market profits, the wholesale marker profits and the accumulated imbalance penalties for each broker are shown in Fig. 9 and Table 4. It can be seen that the retail market profits of COLDPower'16 and maxon16 were similar while the accumulated penalties of maxon16 were less than the accumulated penalties of COLDPower'16. That result signals the importance of studying and including an extra mechanism to minimize the imbalance penalties by using storage units and controllable customers for the balance market.

Finally, it is important to highlight that COLDPower'16 strategy for the wholesale market achieved profits (albeit small), unlike the other brokers that obtained losses in the wholesale market.

7. Conclusions

In this paper, we introduced COLDPower'16, a competitive and profitable multi-agent autonomous broker for energy markets. The different types of energy markets, including retail, wholesale and balance markets were outlined. Power TAC, the energy markets simulator used for reseach comunity, was also described.

The COLDPower'16 modules, algorithms and local strategies for competing in each of the simulated energy markets were presented. The local strategy of each tariff expert agent used reinforcement learning, while the local strategy of the wholesale expert agent estimates future energy prices and the amount of energy that can be negotiated to buy energy when prices are low and sell energy when prices are high. These local strategies contributed to a global strategy that maximizes profit.

The results achieved by COLDPower'16 by in the international tournament Power TAC 2016 were analyzed, showing that COLDPower'16 was both competitive by achieving the 2nd place and also profitable by earning over €86 million EUR.

A comparison with maxon'16, the 1st place broker from Germany, concluded that the future work to make COLDPower'16 even more competitive must include an extra mechanism to minimize imbalance penalties by using storage units and controllable customers in the balance market.

Also, we visualize improving the behavior of COLDPower'16, designing tariff experts that models specific kinds of customers like wind generators, solar farms, and electric vehicles. In addition, for wholesale market, a more complex and effective mechanisms to obtain an estimated projection of future energy prices and the amount of energy that can be traded, like long short-term memory based on deep recurrent neural networks, will be explored.

Acknowledgments

This work is a result of the P12 project of the Mexican Center of Innovation in Energy (CEMIE-Eólico), partially sponsored by Fund (FSE) CONACYT-SENER Energy Sustainability.

References

Buljevic, I., Pranjic, I., Mijic, M., Babic, J., Petric, A., & Podobnik, V. (2012). *The crocodileagent 2012: Reaching agreements in a simulation of a smart grid wholesale market. Proceedings of the first international conference on agreement technologies*. CEUR-W.S.org111–112.

Casazza, J., & Delea, F. (2003). *Understanding electric power systems: An overview of the technology and the marketplace* (2 edition). Wiley-IEEE Press.

Chowdhury, M. M. P., Folk, R. Y., Fioretto, F., Kiekintveld, C., & Yeoh, W. (2017). Investigation of learning strategies for the spot broker in power TAC. In S. Ceppi, E. David, C. Hajaj, V. Robu, & I. A. Vetsikas (Eds.). *Agent-mediated electronic commerce. Designing trading strategies and mechanisms for electronic markets* (pp. 96–111). Cham: Springer International Publishing.

Chowdhury, M. M. P. (2016). *Predicting prices in the power TAC wholesale energy market. Proceedings of the thirtieth AAAI conference on artificial intelligence*. AAAI Press4204–4205.

Diamantopoulos, T. G., Symeonidis, A. L., & Chrysopoulos, A. C. (2013). Designing robust strategies for continuous trading in contemporary power markets. In E. David, C. Kiekintveld, V. Robu, O. Shehory, & S. Stein (Eds.). *Agent-mediated electronic commerce. designing trading strategies and mechanisms for electronic markets* (pp. 30–44). Switzerland: Springer International Publishing.

Fanelli, V., Maddalena, L., & Musti, S. (2016). Asian options pricing in the day-ahead electricity market. *Sustainable Cities and Society, 27*, 196–202.

Golmohamadi, H., Keypour, R., Bak-Jensen, B., & Pillai, J. R. (2019). Optimization of household energy consumption towards day-ahead retail electricity price in home energy management systems. *Sustainable Cities and Society,* 101468.

Hoogland, J., & La Poutré, H. (2017). An effective broker for the power TAC 2014. In S. Ceppi, E. David, C. Hajaj, V. Robu, & I. A. Vetsikas (Eds.). *Agent-mediated electronic commerce. Designing trading strategies and mechanisms for electronic markets* (pp. 66–80). Cham: Springer International Publishing.

Imran, K., & Kockar, I. (2014). A technical comparison of wholesale electricity markets north america and eur-in North America and Europe. *Electric Power Systems Research, 108*, 59–67.

Ketter, W., Collins, J., & Reddy, P. (2013). Power TAC: A competitive economic simulation of the smart grid. *Energy Economics, 39*, 262–270.

Ketter, W., Collins, J., & Weerdt, M.d. (2016). *The 2016 power trading agent competition. Technical Report ERIM Report Series Reference 2016-002-LIS*. Erasmus Research Institute of Management.

Kuate, R. T., He, M., Chli, M., & Wang, H. H. (2013). *An intelligent broker agent for energy trading: an MDP approach. Proceedings of the twenty-third international joint conference on artificial intelligence*. AAAI Press234–240.

Kuate, R. T., Chli, M., & Wang, H. H. (2014). *Optimising market share and profit margin: Smdp-based tariff pricing under the smart grid paradigm. Proceedings of the IEEE PES innovative smart grid technologies Europe*. IEEE1–6.

Lezama, F., Palominos, J., Rodríguez-González, A. Y., Farinelli, A., & Munoz de Cote, E. (2017). Agent-based microgrid scheduling: An ICT perspective. *Mobile Networks and Applications*.

Lezama, F., Soares, J., Hernandez-Leal, P., Kaisers, M., Pinto, T., & Vale, Z. (2018). Local energy markets: Paving the path towards fully transactive energy systems. *IEEE Transactions on Power Systems*.

Liefers, B., Hoogland, J., & La Poutré, H. (2014). A successful broker agent for power TAC. In S. Ceppi, E. David, V. Podobnik, V. Robu, O. Shehory, S. Stein, & I. A. Vetsikas (Eds.). *Agent-mediated electronic commerce. Designing trading strategies and mechanisms for electronic markets* (pp. 99–113). Cham: Springer International Publishing.

Maenhoudt, M., & Deconinck, G. (2010). *Agent-based modelling as a tool for testing electric power market designs. Proceedings of the international conference on the European energy market*. Madrid, Spain: IEEE1–5.

Martinot, E., Chaurey, A., Lew, D., Moreira, J. R., & Wamukonya, N. (2002). Renewable energy markets in developing countries. *Annual Review of Energy and the Environment, 27*, 309–348.

Matetic, S., Babic, J., Matijas, M., Petric, A., & Podobnik, V. (2012). *The crocodileagent 2012: Negotiating agreements in smart grid tariff market. Proceedings of the first international conference on agreement technologies*. CEUR-WS.org203–204.

Mirakhorli, A., & Dong, B. (2018). Market and behavior driven predictive energy management for residential buildings. *Sustainable Cities and Society, 38*, 723–735.

Mont, J. A., & Turner, W. C. (1999). A study on real-time pricing electric tariffs. *Energy Engineering, 96*, 7–34.

Ntagka, E., Chrysopoulos, A., & Mitkas, P. A. (2014). Designing tariffs in a competitive energy market using particle swarm optimization techniques. In S. Ceppi, E. David, V. Podobnik, V. Robu, O. Shehory, S. Stein, & I. A. Vetsikas (Eds.). *Agent-mediated electronic commerce. Designing trading strategies and mechanisms for electronic markets* (pp. 129–143). Cham: Springer International Publishing.

Özdemir, S., & Unland, R. (2018). Agentude17: A genetic algorithm to optimize the parameters of an electricity tariff in a smart grid environment. In Y. Demazeau, B. An, J. Bajo, & A. Fernández-Caballero (Eds.). *Advances in practical applications of agents, multiagent systems, and complexity: The PAAMS collection* (pp. 224–236). Cham: Springer International Publishing.

“Ozdemir, S., & Unland, R. (2015). Trading strategies of a champion agent in a multiagent smart grid simulation platform. In J. P. M”uller, W. Ketter, G. Kaminka, G. Wagner, & N. Bulling (Eds.). *Multiagent system technologies* (pp. 218–232). Cham: Springer International Publishing.

Paiho, S., Saastamoinen, H., Hakkarainen, E., Similä, L., Pasonen, R., Ikäheimo, J., R“am”a, M., Tuovinen, M., & Horsmanheimo, S. (2018). Increasing flexibility of finnish energy systems – A review of potential technologies and means. *Sustainable Cities and Society, 43*, 509–523.

Reddy, [Reddy Veloso, 2011;] P. P., & Veloso, M. M. (2011). *Learned behaviors of multiple autonomous agents in smart grid markets. Proceedings of the twenty-fifth AAAI conference on artificial intelligence*. AAAI Press1396–1401.

Serrano Cuevas, J., Rodriguez-Gonzalez, A. Y., & Munoz de Cote, E. (2017). Fixed-price tariff generation using reinforcement learning. In K. Fujita, Q. Bai, T. Ito, M. Zhang, F. Ren, R. Aydoğan, & R. Hadfi (Eds.). *Modern approaches to agent-based complex automated negotiation* (pp. 121–136). Cham: Springer International Publishing.

Urban, T., & Conen, W. (2017). *Maxon16: A successful power TAC broker. Proceedings of the international conference on autonomous agents and multiagent systems, AAMAS '17*. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems1–14.

Urieli, D., & Stone, P. (2014). *Tactex'13: A champion adaptive power trading agent. Proceedings of the 2014 international conference on autonomous agents and multiagent systems, AAMAS '14*. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems1447–1448.

Urieli, D., & Stone, P. (2016). *An mdp-based winning approach to autonomous power trading: Formalization and empirical analysis. Proceedings of the 2016 international conference on autonomous agents and multiagent systems, AAMAS '16*. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems827–835.

Vargas, A., Saavedra, O. R., Samper, M. E., Rivera, S., & Rodriguez, R. (2016). Latin American energy markets: Investment opportunities in nonconventional renewables. *IEEE Power and Energy Magazine, 14*, 38–47.

Watkins, C. J. C. H., & Dayan, P. (1992). Technical note: q-learning. *Machine Learning, 8*, 279–292.

Wissner, M. (2011). The smart grid-a saucerful of secrets? *Applied Energy, 88*, 2509–2518.

Zou, P., Chen, Q., Yu, Y., Xia, Q., & Kang, C. (2017). Electricity markets evolution with the changing generation mix: An empirical analysis based on China 2050 high renewable energy penetration roadmap. *Applied Energy, 185*, 56–67.