An Analysis of Power TAC 2013 Trial

Jurica Babic and **Vedran Podobnik**
University of Zagreb, Faculty of Electrical Engineering and Computing
Unska 3, HR-10000 Zagreb, Croatia
{jurica.babic,vedran.podobnik}@fer.hr

Abstract

This paper provides insights into performance of competing agents in Power TAC trial held in March 2013. Firstly, the paper gives the description of the Power TAC post-game data set and presents our analysis process. Furthermore, paper discusses the analysis output: indicators about brokers' performance in energy retail market, energy wholesale market as well as the balancing process. Results of the analysis identified diverse approaches in the design of competing agents' strategies, suggesting there will be interesting and successful Power TAC 2013 finals.

Introduction

The current electrical power systems switch from the traditional producer-centric grid to the advanced consumer-centric grid called *smart grid* (Ipakchi and Albuyeh 2009) (Ramchurn et al. 2012). Not only will consumers with installed smart metering equipment be able to adjust their consumption habits according to market price signals received from the smart grid, but also, thanks to new technological solutions, will those consumers become an essential element in real-time alignment of energy demand and supply within the local area. To allow for efficient control of such complex system, retail consumers will be aggregated in the virtual power plants. An example of a planned virtual power plant project is the four-year, €21 million EcoGrid project for the Danish island of Bornholm (Kumagai 2012).

However, in addition to technical aspects of the smart grid, establishment of the supporting market system is crucial. Consequently, what smart grid ecosystem currently lacks in addition to technical infrastructure is an efficient set of market mechanisms. In order to avoid bad market design once smart grids are going to be widely deployed, it is necessary to provide a risk-free environment for testing market policies. The *Power Trading Agent Competition* (Power TAC, *http://powertac.org*) is an open, competitive market simulation platform that addresses the need for policy guidance based on robust research result on the structure and operation of retail electrical energy markets (Ketter, Collins, Reddy, and Weerdt 2012). Power TAC extends the portfolio of TAC games (Collins, Ketter, and Sadeh 2010), "open simulations" that have counterpoised agent-based computational economics (ACE) (Shun-kun and Yuan Jia-hai 2005) as alternative to traditional game-theoretic approaches for testing policies for complex systems (Mohsenian-Rad et al. 2010). The trading agent in Power TAC game is business entity (or "broker") that can fulfill the real-life role of energy retailer in the smart grid environment. Agent's task is to provide energy to consumers through tariff offerings, and then manage its consumer portfolio loads by trading in a wholesale market.

The initial version of Power TAC platform was released in 2011 and has continually been updated with respect to robustness and logical soundness thanks to a lot of support from platform developers as well as the valuable feedback from growing number of teams participating in game trials, which were organized every few months to evaluate the Power TAC platform in the tournament deployment. There are at least three main benefits for using the trial scheme during the Power TAC platform beta phase:

1. Game developers are able to evaluate the platform robustness and identify possible technical issues thanks to intensive trial schedule which contains a variety of game configurations;
2. Multiple approaches in agent's design, coming from independent teams competing in trials, assist in identifying strong and weak points in the game logic implementation. This helps in assuring that the game implementation indeed corresponds to the game specification;
3. Competing teams are able to evaluate the performance of their agents and thus decide what development path

should be taken in terms of agents' features for the final tournament.

Encouraging results from the latest Power TAC 2013 March trial suggest there will be interesting Power TAC 2013 finals, arranged to be held in conjunction with AAAI-13 in July 2013.

This paper describes our approach to Power TAC game analysis and provides insights into performance of each of the competing agents in Power TAC 2013 March trial. The remainder of this paper is organized as follows. Firstly, we provide brief description of the game scenario in Section 2. Afterwards, Power TAC data set used as input for analysis, as well as steps in processing Power TAC data to provide performance indicators, are described in Section 3. Derived performance indicators are presented and discussed in the Section 4. Section 5 concludes the paper with the final remarks about the Power TAC 2013 March trial.

Power TAC

The traditional power systems are defined through the *energy layer* that includes functionalities of energy *production*, *transmission*, *distribution* and *consumption*. The vertical extension of the single-layered traditional power systems with the *information and communications technology (ICT) layer* enables real-time integration of smart grid components and synchronous two-way communication among stakeholders in the power systems. It is believed that the "Internet of energy" (BMWi 2009) will be developed due to use of ICT in energy distribution systems. Consequently, the smart grid ecosystem extends the traditional power grid with various advanced functionalities that are superior to the traditional energy layer functionalities: client-side functionalities are *smart metering* and *demand-side management*, while the grid operator can benefit from *grid balancing* and *real-time monitoring* of the grid. Multi-layered smart grid architecture along with its functionalities and correspondent resource flows is depicted in Figure 1. However, to facilitate the evolution from traditional power systems towards smart grid ecosystem, establishment of the supporting market system is of great importance.

Power TAC is a simulation platform which deals with the *market layer* of the smart grid architecture. The major elements of its scenario are shown in Figure 2. The main element, a competitive trading agent, is a self-interested *broker* that aggregates energy supply and demand with the intent of earning a profit. The majority of broker's energy supply is obtained through the use of *wholesale market*. Brokers must build a good-quality portfolio of *retail customers* (i.e. consumers and producers) by offering carefully designed tariffs through *tariff market*. Good-quality portfolio implies having tariff subscriptions that are profitable and can be real-time balanced. However, the specific consumption and production capacities broker has acquired through the tariff market will almost certainly cause imbalance in broker's energy supply and demand, causing two negative impacts. First, specific broker imbalance contributes to the imbalance of the whole power grid, causing serious problems in the power grid management and lowers the quality of energy provision. The second problem are less-than-attractive balancing fees broker has to pay to distribution utility for causing imbalance of the power grid. Because of described reasons, a profit-oriented broker will tend to use strategies that will contribute to low energy imbalance caused by action in the tariff market. Additionally, to tackle energy imbalance problem, brokers are encouraged to trade in the *wholesale market* by placing bids to acquire some extra energy or to sell an energy excess by placing asks.

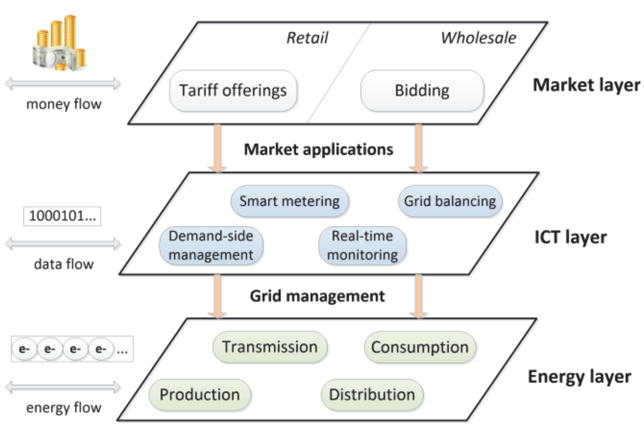

Figure 1: A multi-layered smart grid functional architecture

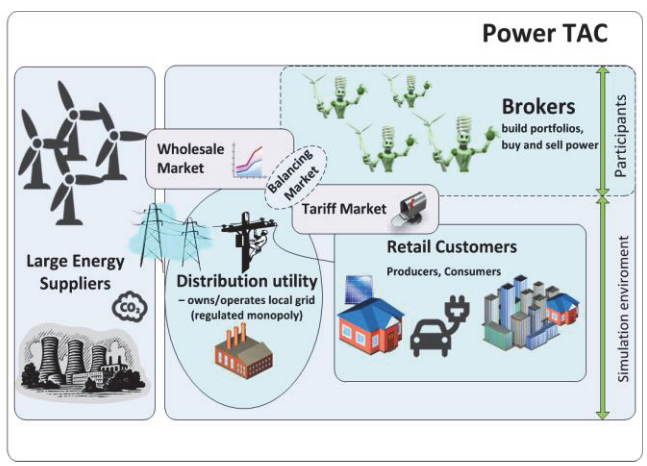

Figure 2: Elements of Power TAC scenario

Methods Used for Analysis of Power TAC Games

In this section we first describe the Power TAC data set as an input for the analysis process. Afterwards, the general idea behind the analysis process is briefly discussed as well.

Data Set

As mentioned in the introduction, we decided to explore the data obtained from Power TAC 2013 March trial. There were five competing teams (hereinafter brokers) participating in the March trial:

- `AstonTAC` (Aston University, UK);
- `CrocodileAgent` (University of Zagreb, Croatia) (Babic et al. 2012) (Buljevic et al. 2012) (Matetic et al. 2012);
- `LARGEPower` (Rotterdam School of Management, Netherlands);
- `Mertacor` (Aristotle University of Thessaloniki, Greece); and
- `MLLBroker` (University of Freiburg, Germany).

The total of 57 games was played through two sets of rounds. In the qualifying round, which had 12 games scheduled, the goal was to ensure brokers communicate with the game server in the correct manner and that they comply with the official rules. Since all brokers completed the first round, the trial tournament proceeded with the final round. The next 45 games were played in the form of three different game sizes: two, three and five players' mode (meaning each game could be played by two, three or five competing teams).

In addition to competing brokers, each Power TAC game contains the embedded agent called `default broker`, which serves in the role of default retailer for customers even before competing brokers join the game. Therefore, we use *gameSize3* (i.e., two competing players and the `default broker`) to denote the Power TAC game with two competing agents, *gameSize4* to denote the Power TAC game with three competing agents and *gameSize6* to denote the Power TAC game with five competing agents.

The output of each Power TAC game is a set of logs for debugging (*trace* log) and game exploration (*state* log) purposes. As the name implies, a *state* log keeps track about all changes to the game state and enables ex-post analysis of brokers, customers and markets.

The chosen input for the analysis process are *state* logs from all 45 games of the final round of Power TAC March trial.

Analysis Process

In order to efficiently perform the analysis, we used the Power TAC *logtool* as the base tool. The existing *logtool* converts an input *state* log in the series of domain objects and thus helps the user to focus only on designing the desired data processing mechanism. Our extended solution of *logtool* parses a defined set of state logs and produces spreadsheets with statistical summaries for the aggregate and individual broker's performances. In particular, the input for analysis presented in this paper are 45 *state* logs from trial's final round while the output are the following artifacts:

- individual game spreadsheet (45 files);
- game size spreadsheet (five files);
- tournament spreadsheet (one file).

Individual game spreadsheet contains the relevant longitudinal data for one game, including overview of broker's transactions in retail, wholesale, balancing and distribution activities. *Game size* and *tournament* spreadsheets contain aggregate data about broker's transactions on the level of different game sizes and the whole tournament, respectively.

Note that there were five g*ame size* spreadsheets, suggesting there were five different game sizes instead of three as described earlier in this paper. The reason for this is the way tournament scheduler works in case there are some problems with one of brokers (i.e. the game will continue without the malfunctioning broker). Careful investigation showed that `Mertacor` broker failed to participate in four games. Even so, those games were included in the trial's tournament level analysis. The discussion on the selected performance indicators is given in the next section.

Power TAC Trial Results & Discussion

Brokers' performance can be observed from various perspectives. The overall performance of the competing brokers is measured through the profit they achieve in the end of the Power TAC game. However, in order to facilitate comparison of brokers' performances in games of different sizes, results are normalized for every game size and the final winner of the tournament is the broker with the highest total (normalized) score. Table 1 shows the official normalized scores of the Power TAC 2013 March trial and reveals that brokers `CrocodileAgent` and `MLLBroker` performed the best.

In the remainder of the section we will explain reasons that stand behind presented final results by examining brokers' performance in several categories: *retail market*, *wholesale market* and *balancing process*.

Table 1: Official normalized scores of the Power TAC 2013 March trial

brokers/ normalized results	gameSize3	gameSize4	gameSize6	Total
CrocodileAgent	0.94	0.85	1.04	**2.83**
MLLBroker	0.68	0.88	1.00	**2.56**
LARGEpower	-0.16	-0.03	0.23	**0.05**
AstonTAC	0.40	0.17	-0.79	**-0.22**
Mertacor	-1.86	-1.86	-1.49	**-5.22**

Retail Market

Retail market is a place where brokers offer their services to retail customers. Since the typical game configuration tends to have much more consumers than producers, brokers are expected to focus on making attractive consumption tariffs which are suitable for various types of consumers. The attractiveness of tariffs offered by various brokers is presented in Figure 3.

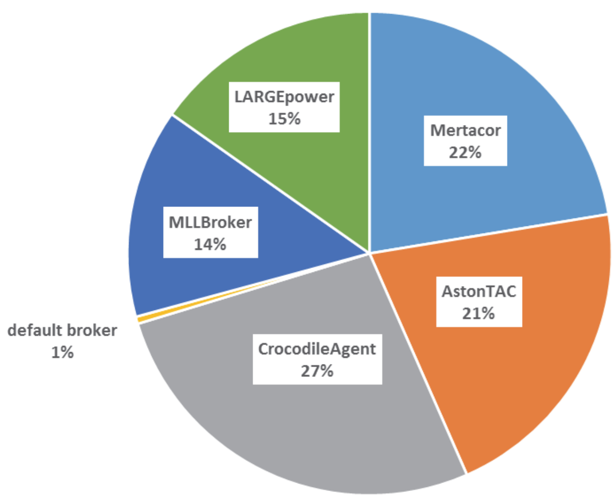

Figure 3: Brokers' energy consumption shares in the retail market

Since `default broker` offers less than attractive consumption tariffs, competing brokers, as expected, did a good job on luring customers away from default tariffs. However, the competing brokers did not enjoy such success with production tariffs, as presented in Figure 4. In case of retail production, only `CrocodileAgent` was able to make reasonable amount of energy bought from tariff transactions along with symbolic 1% share from `MLLBroker`. The reason for this behavior can be examined in some detail using performance indicators from Table 2:

- $\mu_{retailPrice}$ denotes a mean price for tariff transactions. This value is most likely to be negative since the majority of tariff transactions have different signs for energy and cost (i.e. positive energy implies negative cost and vice versa).
- *CR* is a churn rate defined as a ratio between the number of lost customers and the difference between the number of gained and lost customers.
- *Ntariffs* is the total number of published tariff specifications for the whole tournament.
- *TPW* denotes new tariff specifications per week indicator and therefore suggests the level of activity for each broker.
- *WCR* is a wholesale related indicator and it will be explained later in the wholesale market subsection.

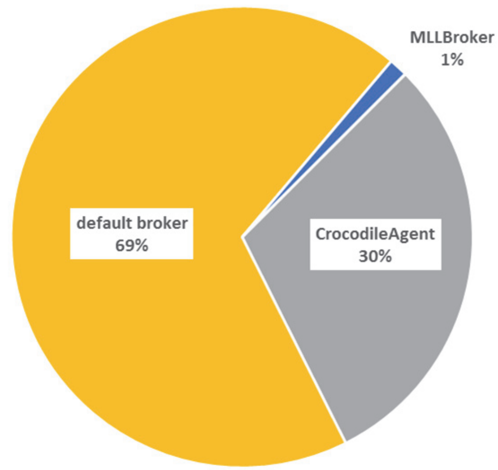

Figure 4: Brokers' energy production shares in the retail market

Table 2: Brokers' retail and wholesale performances

brokers/ performance indicators	$\mu_{retailPrice}$ [€/kWh]	*CR* [%]	$N_{tariffs}$	*TPW*	*WCR*
CrocodileAgent	-0.12	44.7	2 300	10.78	0.95
MLLBroker	-0.05	44.6	100	0.47	0.53
LARGEpower	-0.18	41.5	1 069	4.99	2.67
AstonTAC	-0.45	43.2	25	0.12	0.72
Mertacor	-0.04	45.5	396	3.32	3.11
default broker	-0.26	49.9	90	0.23	1.15
ALL	*-0.13*	*45.6*	*3 980*	*10.36*	*0.79*

The reason for `CrocodileAgent`'s lead in the retail energy department is due to fairly aggressive tariff offering strategy (i.e. over 10 new tariffs each week on average) accompanied with the reasonable unit price of energy. However, the failure to neutralize default broker's production market share is most likely related to the issue with the way the tariffs are being evaluated within the game server. Thus, the results are expected to change once the updated version of the game server will be released.

The trial also proved the high volatile nature of customers' desire to switch over available set of tariffs. Taking into account all customer moves, the average volatility factor defined as a churn rate is 45.6%. Since the brokers did not use signup payments in their tariff offerings, it is expected that the current churn rate will likely decrease once the brokers will use the signup payment option.

Table 3: Retail prices per game size

Brokers/ RMSprice [€/kWh]	*gameSize3*	*gameSize4*	*gameSize6*
`CrocodileAgent`	0.13	0.13	0.13
`MLLBroker`	0.08	0.08	0.08
`default broker`	0.35	0.35	0.31
`LARGEpower`	0.19	0.28	0.25
`AstonTAC`	1.33	1.01	0.23
`Mertacor`	0.06	0.06	0.06
`ALL`	*0.31*	*0.27*	*0.14*

Finally, it is particularly interesting to explore how the root mean square (RMS) retail unit price of energy progresses over game sizes. Judging from values in Table 3, it seems like most of brokers are indifferent to the increase in number of brokers in the market. `AstonTAC` and `LARGEBroker` were only ones with different RMS unit price values across the game sizes. The table also indicates that there is a trend of lowering the unit price when there is more intensive competition.

Wholesale Market

Reliable wholesale price predictions as well as the ability to anticipate the net load from customers are the prerequisites for successful trading in a day-ahead wholesale market[1]. The variety of strategies used can be identified by inspecting brokers' total net volume of traded energy decomposed over the 24-hour interval. Figure 5 reveals individual broker's wholesale strategies. `AstonTAC` make most of its trades reasonably early, predominately 20 to 24 hours ahead. An interesting peak in trading caused by `CrocodileAgent` is noticed between 21 and 20 hours ahead. The period of trade domination for `LARGEpower` is between 19 and 14 hours ahead. `MLLBroker` intensives its net trade volume from 13 to 1 hours ahead. Finally, `Mertacor` is able to make trades exclusively in the last two hours (i.e. 2 or 1 hours ahead). The period between 21 and 12 hours ahead is where `default broker` has a negative net value of traded energy on wholesale market. This implies that the aforementioned broker was selling energy in that interval, most likely due to the strategy of buying more energy than needed.

Additional remarks regarding bidding strategies are made with the help of wholesale clearing rate indicator *WCR* in Table 2, defined as the ratio between the number of successful trades and the number of submitted orders. If *WCR* indicator is close to zero then the observed broker frequently submits orders that will not clear on the market.

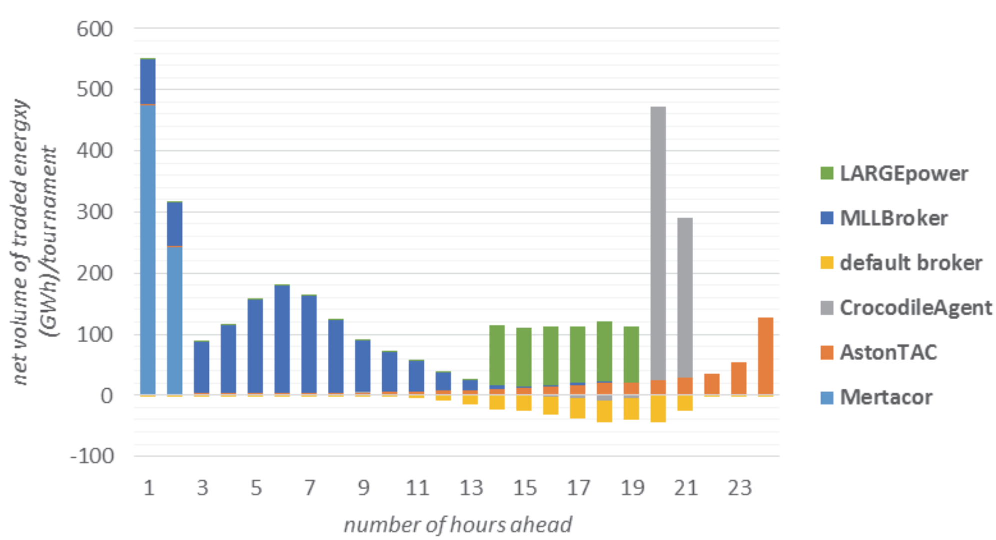

Figure 5: Net volume of traded energy decomposed by number of hours ahead

[1] In the day-ahead market, contracts are made between seller and buyer for the delivery of power in the next 24 hours (i.e. the price is set and the trade is agreed).

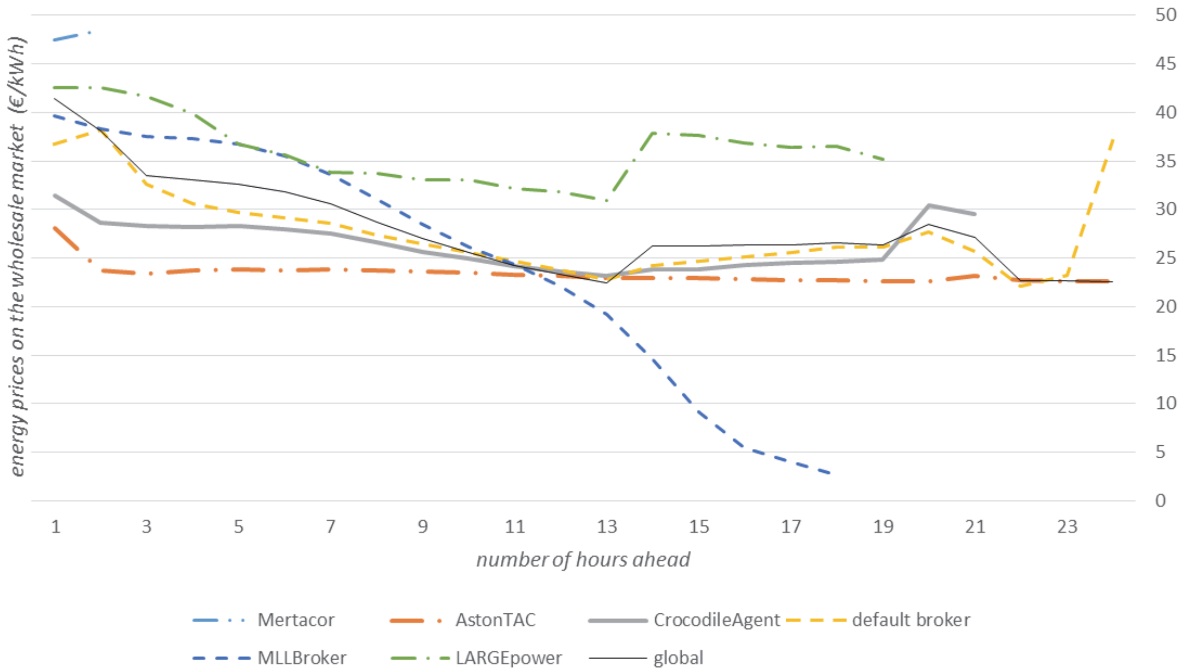

Figure 6: RMS wholesale prices for the day-ahead market

If the *WCR* is too high[2], then the broker frequently sends offers that are accepted by other participants, meaning the broker most likely could have made a better deal. In that case, `Mertacor` and `LARGEPower` used over-compatible orders while `MLLBroker` had the lowest percentage of its cleared orders.

The mean wholesale prices, which vary over 24-hour timeframe, are presented in Figure 6. The general trend for mean wholesale prices proved to be increasing towards the last possible hour ahead. The lowest mean price (22.58 €/MWh) was recorded in the first possible hour ahead of trading. It is also worth mentioning that `CrocodileAgent` follows the global mean wholesale price and it even manages to score some trades with cheaper than global mean price in the last hours ahead. An interesting period of cheap trading for `MLLBroker` is between 18 and 12 hours ahead. `LARGEPower` mean price suggest it consistently makes trades at slightly higher prices. `AstonTAC` keeps its mean price within tight boundaries while `Mertacor` only trades in the final two hours at highest prices.

Balancing

Balancing process occurs in the last phase of each hour[3] of the game. Brokers that do not have balanced customer portfolios and do not have successful wholesale bidding strategy mechanisms will be charged with the less than attractive balancing cost.

In order to examine the balancing performance of each broker, a set of imbalance indicators were prepared in Table 4:

- $RMS_{imbalance}$ is a root mean square imbalance for balancing transactions. Since the imbalance can be negative (i.e. broker has an energy deficit) or positive (i.e. broker has an energy surplus), we opted to use RMS which uses squared i.e. positive values.
- $\mu_{imbalance}$ is a mean imbalance in kilowatts.
- *IR* indicator is defined as the ratio of sum of energy imbalance and sum of energy consumption.
- $\mu_{contribution}$ tells how much did the broker contribute to the total imbalance. A higher value suggests the broker was on average more imbalanced than the other broker. A negative value (not the case in this trial) suggests broker on average had imbalance of a different sign than the total imbalance. If the value is close to zero than the broker performed well during the balancing process.
- μ_{price} is a mean unit balancing price. A positive value means the broker pays to the grid operator because of broker's energy deficit. Similarly, a negative value means the grid operator pays to the broker because of broker's energy surplus.

[2] Please note *WCR* can also be larger than 1 because each submitted order can have multiple successful transactions (i.e., from multiple market players)

[3] A simulated hour is discrete time unit in the Power TAC game.

Table 4: Imbalance stats

Imbalance stats	$RMS_{imbalance}$ [kWh]	$\mu_{imbalance}$ [kWh]	IR [%]	$\mu_{contribution}$ [kWh]	μ_{price} [€/kWh]
`CrocodileAgent`	10 571.93	-6 253.64	22	3 261.26	0.04
`MLLBroker`	28 304.62	16 426.29	-109	15 502.93	-0.03
`LARGEpower`	3 908.20	-382.39	2	804.12	0.09
`AstonTAC`	22 028.01	-11 084.68	49	11 247.26	0.08
`Mertacor`	39 162.03	-31 093.93	72	31 894.70	0.09
`default broker`	3 150.29	-181.70	59	250.14	0.09

Thanks to decent energy consumption share and last-hour wholesale trading, `Mertacor` had on average the highest imbalance charge. `Mertacor`'s *IR* indicator suggests as much as 72% of its energy consumption had to be provided by the balancing utility. The best balanced broker is `LARGEPower`, only 2% of its consumption was left imbalanced. `CrocodileAgent` is reasonably well balanced due to the fact that it is a leader in the energy consumption and still scores reasonably good (i.e., 0.04 €/kWh) mean balancing price. Mean imbalances suggest all brokers had deficit of energy except `MLLBroker` who was on average on the energy surplus. The reason for this lies in the fact `MLLBroker` was consistently buying more energy than its customers needed. The proof of this is its extreme IR indicator and negative μ_{price}.

Dependencies among discussed performance on wholesale market, retail market and balancing process can be seen in Figure 7. All brokers achieved positive retail balance (i.e. sum of all tariff transactions, both with consumers and producers) meaning there were actually making profit on the retail market. This is expected since the majority of retail population is consumption-driven, meaning most of customers pay to brokers for using the electricity.

It is also interesting to see how the majority of `Mertacor`'s losses are due to high penalties exercised by the balancing process. In contrast, `AstonTAC` who has similar balance on the retail market scored much better during the balancing process. However, the low wholesale cost `AstonTAC` experienced during the trial suggests its early bidding strategy did not pay off in terms of acquiring the proper amount of energy.

`LARGEPower` proved its dominance in the balancing process by having roughly the same balancing cost as `default broker`. However, `LARGEPower` performs much better than `default broker` because the latter had much less energy distributed, as witnessed by distribution cost in both brokers.

`CrocodileAgent` was able to win the trial over `MLLBroker` thanks to better wholesale performance and consequently lower balancing cost. A quick observation of retail balance and distribution costs from these brokers confirms `MLLBroker` had a larger positive retail balance and lower distribution cost than `CrocodileAgent`, meaning `MLLBroker` has managed to charge higher margin for its services than `CrocodileAgent`. This all lead to a clue that `MLLBroker` had somewhat more optimized retail strategy than `CrocodileAgent`.

Conclusion

The Power TAC 2013 March trial has proven that the Power TAC simulation platform produces analyzable results. The five competing teams confronted their brokers in both qualifying and final rounds. We analyzed performances of competing brokers by taking into consideration all *state* logs from the final round.

Different brokers implemented distinct behaviors in retail and wholesale markets, as well as during balancing process. Judging from the presented analysis,

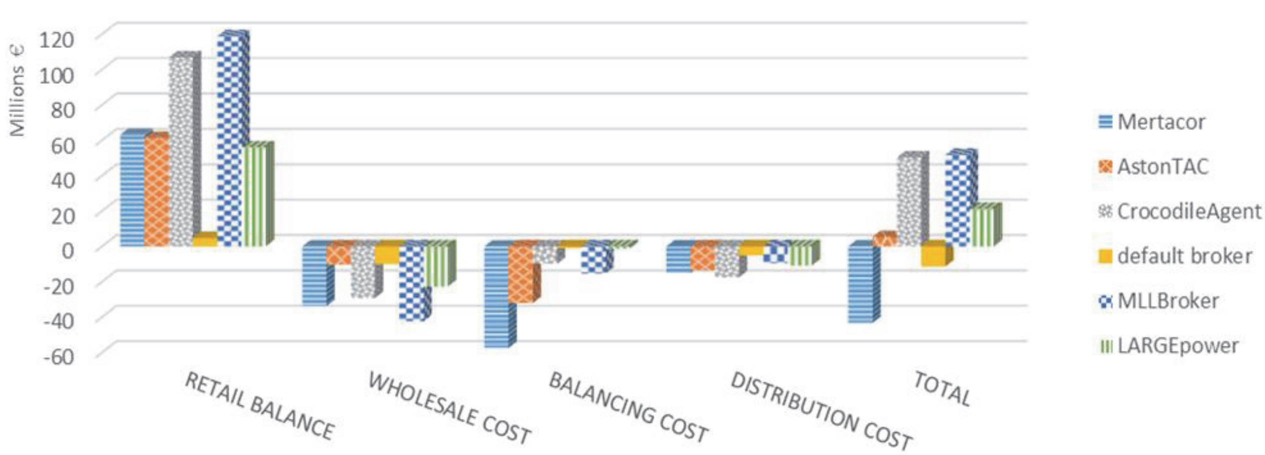

Figure 7: Brokers' cost profiles by categories

CrocodileAgent topped the retail energy consumption share thanks to its aggressive tariff offerings. The award for most interesting wholesale behavior goes to MLLBroker thanks to its extremely cheap wholesale mean prices for 18 to 12 hours ahead. The best balancing broker was LARGEPower because only 2% of its energy consumption was left imbalanced.

Encouraging results from the analyzed trial suggest we will witness interesting and successful Power TAC finals, arranged to be held in conjunction with AAAI-13 in July 2013.

Acknowledgments

The authors acknowledge the support of research project "Intelligent Software Agents for Managing Processes on Electronic Markets", supported by the Foundation of the Croatian Academy of Sciences and Arts, as well as research project "Content Delivery and Mobility of Users and Services in New Generation Networks", funded by the Ministry of Science, Education and Sports of the Republic of Croatia.

Furthermore, authors thank Power TAC game master for organizing Power TAC 2013 March trial as well as competing teams for participation.

References

Babic, J.; Matetic, S.; Matijas, M.; Buljevic, I.; Pranjic, I.; Mijic, M.; and Augustinovic, M. 2012. The CrocodileAgent 2012: Research for Efficient Agent-based Electricity Trading Mechanisms.In *Proceedings of Special Session on Trading Agent Competition @ 6th International KES Conference on Agents and Multiagent Systems – Technologies and Applications (KES-AMSTA 2012)*, 1-13.

Buljevic, I.; Pranjic, I.; Mijic, M.; Babic, J.; Petric, A.; and Podobnik, V. 2012. The CrocodileAgent 2012: Reaching Agreements in a Simulation of a Smart Grid Wholesale Market. In *Proceedings of the 1st International Conference on Agreement Technologies (AT 2012)*, 111-112.

BMWi. 2009. *E-Energy – IKT-basiertes Energiesystem der Zukunft*. Stand.

Collins, J.; Ketter, W.; and Sadeh, N. 2010. Pushing the limits of rational agents: The trading agent competition for supply chain management. *AI Magazine* 31(2):63–80.

Ipakchi, A.; and Albuyeh, F. 2009. Grid of the future. *Power and Energy Magazine. IEEE*. 7(2):52-62.

Ketter, W., Collins, J., Reddy, P. P., and Weerdt, M. D. 2012. The 2012 Power Trading Agent Competition. *ERIM Report Series Reference No. ERS-2012-010-LIS.*

Kumagai, J. 2012. Virtual Power Plants, Real Power. *IEEE Spectrum* 49(3): 13-14.

Matetic, S.; Babic, J.; Matijas, M.; Petric, A.; and Podobnik, V. 2012. The CrocodileAgent 2012: Negotiating Agreements in a Smart Grid Tariff Market. In *Proceedings of the 1st International Conference on Agreement Technologies (AT 2012)*, 203-204.

Mohsenian-Rad, A.-H.; Wong, V.W.S.; Jatskevich, J.; Schober, R.; and Leon-Garcia, A. 2010. Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid. IEEE Transactions on Smart Grid 1(3): 320-331.

Ramchurn, S.D.; Vytelingum, P.; Rogers, A.; Jennings, N.R. 2012. Putting the 'smarts' into the smart grid: a grand challenge for artificial intelligence. Communications of the ACM 55(4): 86-97.

Shun-kun, Y. and Jia-hai, Y. 2005. Agent-based computational economics: methodology and its application in electricity market research. In *7th International Power Engineering Conference (IPEC 2005)*, 1-5.