VidyutVanika: An Autonmous Broker Agent for Smart Grid Environment

Susobhan Ghosh[1], Kritika Prakash[1], Sanjay Chandlekar[2]

Easwar Subramanian[2], Sanjay P. Bhat[2], Sujit Gujar[1] and Praveen Paruchuri[1]

[1] IIIT Hyderabad, India {susobhan.ghosh, kritika.prakash}@research.iiit.ac.in, {sujit.gujar, praveen.p}@iiit.ac.in
[2] TCS Innovation Labs, Hyderabad, India {sanjay.rc, easwar.subramanian, sanjay.bhat}@tcs.com

Abstract

We describe the design of an autonomous electricity broker agent, VidyutVanika, the runner-up of the 2018 PowerTAC competition. The agent uses techniques from reinforcement learning, dynamic programming and other areas of machine learning to seek appropriate actions in tariff and wholesale market of the PowerTAC simulation environment. The novelty of our agent lies in defining the reward functions of suitably defined Markov decision processes (MDPs), solving these MDPs, and applying these solutions to real actions in the market. In addition, VidyutVanika uses a neural network to predict the energy consumption of various customers using weather data. The usage forecasts, so obtained, are used to place orders in day-ahead wholesale market. These forecasts also helps in reducing the balancing costs incurred by the broker.

Introduction

A *smart grid* is an electrical grid built on an sophisticated infrastructure to manage electricity demand in a sustainable, reliable and economical manner using smart meters, smart appliances and renewable energy sources. An integral component of a smart gird is the electricity distributing agencies or *brokers* who serve retail customers by buying energy in bulk from generating companies. Brokers handle supply-demand imbalance through dynamic pricing strategies, suitable use of storage devices and tapping renewable energy from small-time producers. There are multiple challenges in the realization of smart grids, like managing highly fluctuating supply-demand scenarios, engaging stakeholders with ulterior motives, and handling automation failures of participating entities (Ketter et al. 2016b; 2016a). In order to foresee such problems and examine potential solutions, PowerTAC (Ketter, Collins, and Weerdt 2017) provides an open source simulator platform that replicates crucial elements of a smart grid system and allows large-scale experimentation. The simulation encourages the development of autonomous broker agents that aim at making a profit by offering electricity tariffs to customers in a retail (or tariff) market, and trading energy in a competitive wholesale market, while carefully balancing their supply and demand. To this end, a Power Trading Agent Competition (Power TAC) (Ketter, Collins, and Weerdt 2017) is held annually.

Since 2012, several research groups have benchmarked, deployed and published strategies using PowerTAC. Popular teams include AgentUDE (Özdemir and Unland 2018a) (2015; 2018a; 2018b), TacTex (Urieli and Stone 2014) (2016), SPOT (Chowdhury et al. 2017; 2018; Chowdhury 2016) and Maxon (Urban and Conen 2017) to name a few. These studies have demonstrated that machine learning and game theory-based strategies are essential for such broker agents to dynamically price tariffs and predict customer usage while simultaneously placing bids in wholesale auctions. Broker agents have also used Markov Decision Process (MDP) to model strategies in the tariff market (Cuevas, Rodriguez-Gonzalez, and De Cote 2017), and wholesale market (Urieli and Stone 2014; 2016; Reddy and Veloso 2011), while others have employed genetic algorithm, fuzzy-logic and tailored heuristics for the same (Özdemir and Unland 2018a; Rúbio et al. 2015; Liefers, Hoogland, and La Poutré 2014).

Our goal is to design a learning broker with the following objectives: (i) React to competing tariffs (ii) Increase market share, i.e., subscribed customers (iii) Decrease transmission capacity costs (iv) Decrease costs of energy procurement. Taking cue from (Urieli and Stone 2014; Reddy and Veloso 2011; Cuevas, Rodriguez-Gonzalez, and De Cote 2017) we formulate our tariff and wholesale market problems as separate MDPs. Though our MDP formulations are motivated by (Urieli and Stone 2014) and (Cuevas, Rodriguez-Gonzalez, and De Cote 2017), our novelty lies in their reward structure, solution, and application of those solutions. These are supplemented by a neural network based usage predictor, that also utilizes weather data. Our broker, VidyutVanika, referred to as VV throughout the paper, was the runner-up in PowerTAC 2018 Finals. In this note, we describe the various aspects of our broker and showcase its efficacy with the help of data gathered from PowerTAC 2018 competition. In the end, we also describe the ongoing improvements to the broker design for the 2019 edition of PowerTAC tournament. We note that a detailed version of this short exposition has appeared elsewhere (Ghosh et al. 2019).

Overview of Broker Agent

This section presents an overview of our broker agent, VV. The broker consists of two main modules, namely, *Tariff Module* (TM) and *Wholesale Module* (WM). TM is responsible for publishing and revoking tariffs in the tariff (or retail) market. WM generates bids/asks to purchase/sell energy contracts in the wholesale market. VV doesn't actively participate in the balancing market. Tariff design is accomplished by formulating a MDP (Puterman 1994), which we approximately solve using Q-learning (Watkins and Dayan 1992). We model the bidding problem in the wholesale market as a separate MDP, which we solve using dynamic programming (Bellman 2013). In addition to these two modules, VV incorporates a Customer Usage Predictor (CUP) sub-module built using neural networks (NN) to predict the usage of all subscribed customers in a future time slot, by using weather forecasts and past usage pattern of each customer. VV aggregates the predicted usage across all its subscribed customers to estimate the amount of energy to be procured in the wholesale market. Doing so helps VV reduce the imbalance on its portfolio.

Tariff Module (TM)

The tariff module of VV, maintains two active time-of-use (TOU) tariffs, namely, (i) MDPTOU and (ii) WeeklyTOU. MDPTOU is the result of solving an MDP problem for retail market using Q-learning, and is revised every twenty-four hours. WeeklyTOU is an empirically determined, fixed weekly TOU tariff, which remains active throughout the duration of the game.

Generating MDPTOU is a two-step process - (1) Generate a Fixed Price Tariff (FPT) by solving an MDP using Q-learning; (2) Convert the FPT to a TOU tariff for consumption customers by predicting the overall demand profile for the tariff market over the next 24 time slots.

Our Tariff MDP formulation is primarily motivated from the work of (Cuevas, Rodriguez-Gonzalez, and De Cote 2017). At any simulation time t, the state of the MDP is a quadruple that captures four features of the tariff market. The first feature is rationality of the tariff market which is decided based on whether the highest production tariff is lower or higher than the lowest consumption tariff. The second is the portfolio status of our broker agent VV which could be surplus, balanced or deficit depending on the difference between the amount of energy acquired and committed in the tariff market at time t. The third and fourth features rank the VV's current consumption and production tariffs with respect to prevailing tariffs of other competing broker agents. In total, there are 96 possible states in the MDP. The action set is discrete and consists of 8 actions, each of which lets VV modify its previous production and consumption tariff in a specific fashion. These include different combinations of increasing, decreasing or tempering the latest prevailing production or consumption tariff of broker VV. A detailed description of the state and action space of the MDP can be found in (Cuevas, Rodriguez-Gonzalez, and De Cote 2017).

The key novelty in our MDP formulation is the reward structure c.f. Cuevas, Rodriguez-Gonzalez, and De Cote. The idea behind the reward structure is to capture the net profit made by VV when it incurs no balancing charge. Thus, the reward at time t is given by:

$$r_t = \theta_{t,C} P_{t,C} - \theta_{t,P} P_{t,P} - \theta_{t,W} W_t \tag{1}$$

The first term in Equation 1 represents the revenue generated by selling energy $\theta_{t,C}$ at the tariff $P_{t,C}$ to consumers of VV at time t. Similarly, the second term represents the amount paid to producers of VV for procuring energy $\theta_{t,P}$ at the tariff $P_{t,P}$. The third term represents the amount paid in the wholesale market to satisfy the net unfulfilled demand $\theta_{t,W} = \theta_{t,C} - \theta_{t,P}$ at unit wholesale procurement cost W_t.

The aforementioned MDP is solved using Q-learning (Watkins and Dayan 1992). Specifically, we construct a Q-table for possible state-action pairs using suitable discount rate γ by playing several offline games with multiple brokers as opponents. The Q-table thus learnt, is then used to arrive at suitable tariffs, while playing games in the real tournament.

Whereas, for production customers, the tariff suggested by the MDP agent is published without any change, for consumption customers, the FPT tariff is converted to a TOU tariff before being published. To this end, the agent first predicts the net demand in the PowerTAC market for the next twenty-four hours of the simulation. Thereafter, at each of the next twenty-four timeslots, the FPT is modified by an amount that is proportional to the excess estimated net demand at that timeslot over the mean estimated demand for the twenty-four hour period. For details, the reader is referred to (Ghosh et al. 2019). The TOU tariffs, thus published, helps in offsetting some of the peak demand charges.

Wholesale Module (WM)

In order to balance the future net usage in its tariff portfolio, VV participates in the wholesale market auctions by placing bids/asks of the form (*energy amount, limit-price*). Neural networks (NN) are used to predict the net usage of a future timeslot. A MDP is formulated and solved using dynamic programming to determine a suitable limit-price for auctions in wholesale market. Further, VV procures the predicted net usage for a timeslot $t + 24$ by participating in twenty-four possible auctions from $\{t, \ldots, t + 23\}$. This is done with the aim of buying more and selling less in those auctions in which the prices are expected to be low, and vice-versa.

The customer usage prediction module (CUP) is responsible for predicting the net usage of the broker's tariff portfolio for a future target time-slot t, by summing over the predicted usage of each customer subscribed to the broker for that target time-slot t. To this end, for each customer, we deploy a small feed forward neural net with input data consisting of the actual weather data, time of day (0-23), and day of week (1-7), while the target variable is the actual usage of the customer. During prediction, the weather forecast is used in place of the actual weather data to predict the usage for the next 24 hours. The model is improved as more data points become available during the game.

VV's Limit Price Predictor is primarily motivated by the work of (Urieli and Stone 2014) on MDP-based wholesale bidding strategy, which in turn is based on (Tesauro and Bredin 2002). Although we use a similar MDP structure, the novelty lies in the reward, solution and application to place bids. First, we do not the bid for entire predicted energy requirement in a single auction as proposed by Urieli and Stone. Rather we participate in twenty-four possible auctions to procure the required amount of energy for a future time slot with the aim of buying more and selling less in those auctions in which the prices are expected to be low, and vice-versa. Second, we use the limit-prices obtained by solving the MDP to place several small bids to purchase small quantities of energy. These small bids help in calculating better estimates for the probability of a bid getting cleared for a given limit price.

VV maintains two instances of the MDP at all times - one for bids, another for asks. The state of the wholesale MDP is the number of bidding opportunities left to buy energy for a future time slot. The action is a limit price that would be used in the bidding process. The reward is the amount of cost incurred in obtaining the total amount of energy required for a future time-slot. The detailed description of the MDP can be found in (Ghosh et al. 2019). The solution to the MDP is a sequential bidding strategy that minimizes the cost per unit energy procured. It is given by a value function which equals the *balancing-price* at the terminal state and is recursively solved at other states using dynamic programming. The solution gives an optimal *limit-price* for each state auction.

Results

The Power TAC 2018 Finals had 7 brokers from research groups across the globe. The tournament had a total of 324 games, with all possible combinations of 7-broker games (100 games), 4-broker games (140 games; 80 games for each broker), and 2-broker games (84 games; 24 games for each broker). Table 1 shows the net profit of all brokers across different game configurations, percentage of profit in comparison to the winning agent, AgentUDE, and the corresponding normalized scores. Despite winning more games than AgentUDE, VV was placed next to AgentUDE in overall ranking of Power TAC 2018. This is because, the determination of the winner is made based on normalized cumulative profits in each configuration across all games in the tournament. Specifically, AgentUDE netted high profits against competing agents (excluding VV) in 2-player games that helped in cementing its place as the winner of the tournament.

Table 2 shows the number of 1^{st} and 2^{nd} place finishes by each broker across all three configurations. As seen, VV won the most number of games in the tournament with 112 wins out of the 204 it participated in, with AgentUDE coming second with 92 wins out of 204. VV had the most wins in 7-broker and 4-broker games, and had the second highest number of wins, behind AgentUDE, in 2-broker games. It is important to note that, overall, VV finished in the top two, 72% of the time whenever it played in a game with more than 2 brokers. In comparison, AgentUDE stood at 65%. On a head-to-head comparison with AgentUDE, out of 100 7-broker games, AgentUDE and VV both shared 39 wins each. However in 4-Broker games in which both VV and AgentUDE participated, VV won 31 times out 40, with AgentUDE winning the remaining 9. In the four 2-broker games involving both brokers, AgentUDE ended up winning three games. VV led in all these three lost games almost till the end, only to fall behind finally due to transmission capacity fees. We also looked at the number of games in which each broker ended up with a negative profit. CrocodileAgent had the fewest games with negative profits, with VV coming second in this category with four times the average market share. Thus, VV managed to make up for its losses on a consistent basis, and rarely ended up being non-profitable.

The tariff module played a crucial role in VV's success, offering tariffs which were attractive to majority of the customers and contributed the most in revenue. VV had the highest market share on average in 2-broker games, 7-broker games and overall, and the second highest in 4-broker games. In contrast, AgentUDE had only a quarter of the overall average market share of VV. While one may expect a greater market share to lead to more profits, it usually leads to higher transmission capacity fees and distribution costs, which can cause higher losses unless managed properly. As a result, agents with lower market share often tend to make less losses, and end up winning. VV also had one of the best tariff market income-to-cost ratio (1.14), with only AgentUDE (1.43) and CrocodileAgent (1.32) having better ratios. However, both AgentUDE and CrocodileAgent had very low average market share compared to VV. Thus, VV is very efficient at making profits despite having a higher market share. Finally, although there was no explicit strategy for balancing market, VV had less imbalance costs even with high market share which exhibits the effectiveness of net usage prediction strategy using neural networks.

Conclusion

We described the critical elements of the strategy used by our broker VidyutVanika (VV), the runner-up in Power TAC 2018 Finals. In particular, we described details of our two modules, TM and WM. TM and WM were responsible for VidyutVanika's actions in the tariff and wholesale market, respectively. The novelty of VidyutVanika lay in (i) defining reward functions for the MDPs, (ii) solving the MDPs, (iii) applying the MDP solutions to actions in the markets, and (iv) NN based usage predictor incorporating available weather data for better customer usage prediction.

Future work

We wish to bring in the following improvements to VV in the 2019 version of PowerTAC competition. The customer usage predictor will contain sophisticated neural nets like LSTM to be able to do p-step ahead prediction of the usage. The tariff module is being designed to roll out price-based demand response (PBDR) tariffs (Valogianni and Ketter 2016) to mitigate the effects of capacity transaction charges. Other changes that are being contemplated include a clever use of storage and electric vehicle customers (Kahlen, Ketter, and van Dalen 2018) in balancing market. The effectiveness of these modules in a multi-agent setting will be studied during the 2019 PowerTAC tournament.

References

Bellman, R. 2013. *Dynamic programming*. Courier Corporation.

Chowdhury, M. M. P.; Folk, R. Y.; Fioretto, F.; Kiekintveld, C.; and Yeoh, W. 2017. Investigation of learning strate-

Broker	7-broker	4-broker	2-broker	Total	7-broker (N)	4-broker (N)	2-broker (N)	Total (N)
AgentUDE	49964603 (100)	62138484 (100)	134908672 (100)	247011760 (100)	1.091	0.634	1.565	3.291
VidyutVanika	48197051 (96)	101942819 (164)	47541635 (35)	197681504 (80)	1.056	1.061	0.336	2.453
CrocodileAgent	27659543 (55)	45441732 (73)	62881837 (47)	135983111 (55)	0.648	0.455	0.552	1.655
SPOT	-6979768 (-14)	32981756 (53)	49183707 (36)	75185695 (30)	-0.041	0.322	0.359	0.64
COLDPower18	2063729 (4)	10289982 (17)	521330 (0.3)	12875040 (5)	0.139	0.078	-0.326	-0.109
Bunnie	-67983216 (-136)	-25049555 (-40)	-19596577 (-15)	-112629348 (-46)	-1.254	-0.3	-0.609	-2.163
EWIIS3	-87271195 (-175)	-206960249 (-333)	-109800161 (-81)	-404031605 (-164)	-1.638	-2.25	-1.878	-5.766

Table 1: Power TAC 2018 – Net profits and normalized scores (denoted by (N)) of each broker

Brokers	7-Broker		4-Broker		2-Broker		Total	
	1^{st}	2^{nd}	1^{st}	2^{nd}	1^{st}	2^{nd}	1^{st}	2^{nd}
VidyutVanika	39	21	54	14	19	5	55	20
AgentUDE	39	26	31	21	22	2	45	24
CrocodileAgent	8	34	13	41	15	9	18	41
SPOT	0	0	16	19	9	15	12	17
COLDPower18	0	3	5	29	8	16	6	24
Bunnie	13	15	21	16	9	15	21	22
EWIIS3	1	1	0	0	2	22	1	11

Table 2: Power Tac 2018 – Number of 1^{st} and 2^{nd} place standings of each broker

gies for the SPOT broker in Power TAC. In Ceppi, S.; David, E.; Hajaj, C.; Robu, V.; and Vetsikas, I. A., eds., *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*, 96–111. Cham: Springer International Publishing.

Chowdhury, M. M. P.; Kiekintveld, C.; Son, T. C.; and Yeoh, W. 2018. Bidding strategy for periodic double auctions using Monte Carlo tree search. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, 1897–1899. International Foundation for Autonomous Agents and Multiagent Systems.

Chowdhury, M. M. P. 2016. Predicting prices in the Power TAC wholesale energy market. In *AAAI*, 4204–4205.

Cuevas, J. S.; Rodriguez-Gonzalez, A. Y.; and De Cote, E. M. 2017. Fixed-price tariff generation using reinforcement learning. In *Modern Approaches to Agent-based Complex Automated Negotiation*. Springer. 121–136.

Ghosh, S.; Subramanian, E.; Bhat, S. P.; Gujar, S.; and Paruchuri, P. 2019. Vidyutvanika: A reinforcement learning based broker agent for a power trading competition. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, 2019*, AAAI 2019.

Kahlen, M. T.; Ketter, W.; and van Dalen, J. 2018. Electric vehicle virtual power plant dilemma: Grid balancing versus customer mobility. *Production and Operations Management* 27(11):2054–2070.

Ketter, W.; ; Peters, M.; Collins, J.; and Gupta, A. 2016a. Competitive benchmarking: An IS research approach to address wicked problems with big data and analytics. *MIS Quarterly* 40(4):1057–1080.

Ketter, W.; ; Peters, M.; Collins, J.; and Gupta, A. 2016b. A multiagent competitive gaming platform to address societal challenges. *MIS Quarterly* 40(2):447–460.

Ketter, W.; Collins, J.; and Weerdt, M. 2017. The 2018 power trading agent competition.

Liefers, B.; Hoogland, J.; and La Poutré, H. 2014. A successful broker agent for Power TAC. In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*. Springer. 99–113.

Özdemir, S., and Unland, R. 2015. Autonomous power trading approaches of a winner broker. In *AAMAS Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis (AMEC/TADA 2015)*, 143–156. Springer.

Özdemir, S., and Unland, R. 2018a. AgentUDE17: A genetic algorithm to optimize the parameters of an electricity tariff in a smart grid environment. In *Advances in Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection*, 224–236. Springer.

Özdemir, S., and Unland, R. 2018b. AgentUDE17: Imbalance management of a retailer agent to exploit balancing market incentives in a smart grid ecosystem. In *Multikonferenz Wirtschaftsinformatik MKWI, 2018.*

Puterman, M. L. 1994. Markov decision processes. *Wiley and Sons.*

Reddy, P. P., and Veloso, M. M. 2011. Strategy learning for autonomous agents in smart grid markets. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Two*, IJCAI'11, 1446–1451. AAAI Press.

Rúbio, T. R.; Queiroz, J.; Cardoso, H. L.; Rocha, A. P.; and Oliveira, E. 2015. TugaTAC broker: A fuzzy logic adaptive reasoning agent for energy trading. In *Multi-Agent Systems and Agreement Technologies*. Springer. 188–202.

Tesauro, G., and Bredin, J. L. 2002. Strategic sequential bidding in auctions using dynamic programming. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 2*, AAMAS '02, 591–598. New York, NY, USA: ACM.

Urban, T., and Conen, W. 2017. Maxon16: A successful Power TAC broker. In *International Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis (AMEC/TADA 2017).*

Urieli, D., and Stone, P. 2014. TacTex'13: A champion adaptive power trading agent. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 465–471. Association for the Advancement of Artificial Intelligence.

Urieli, D., and Stone, P. 2016. Autonomous electricity trading using time-of-use tariffs in a competitive market. In *Pro-*

ceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16). Association for the Advancement of Artificial Intelligence.

Valogianni, K., and Ketter, W. 2016. Effective demand response for smart grids: Evidence from a real-world pilot. *Decision Support Systems* 91:48–66.

Watkins, C. J., and Dayan, P. 1992. Q-learning. *Machine learning* 8(3-4):279–292.