TacTex'13: A Champion Adaptive Power Trading Agent (Extended Abstract)

Daniel Urieli
Department of Computer Science
The University of Texas at Austin
urieli@cs.utexas.edu

Peter Stone
Department of Computer Science
The University of Texas at Austin
pstone@cs.utexas.edu

ABSTRACT

Sustainable energy systems of the future could no longer rely on the current paradigm that energy supply follows demand. Since many of the renewable energy resources do not produce power on demand, there is a need for new market structures that motivate sustainable behaviors by participants. The Power Trading Agent Competition (Power TAC) is a new annual competition that focuses on the design and operation of future retail power markets, specifically in smart grid environments with renewable energy production, smart metering, and autonomous agents acting on behalf of customers and retailers. Its purpose is to help researchers understand the dynamics of customer and retailer decision-making, as well as the robustness of proposed market designs. This research contributes to the former, by introducing TacTex'13, the champion agent from the inaugural competition in 2013. TacTex'13 learns and adapts to the environment in which it operates, by heavily relying on reinforcement-learning and prediction methods. We formalize the complex decision-making problem that TacTex'13 faces, and approximate its solution in TacTex'13's constituent components. We examine the success of the complete agent through analysis of competition results.

Categories and Subject Descriptors

I.2 [**Computing Methodologies**]: Artificial Intelligence

Keywords

Machine Learning, Energy Trading, Smart-Grid

1. INTRODUCTION

Sustainable energy systems of the future will have to include a robust solution to a major challenge presented by many of the renewable energy resources (wind, solar, tidal, etc.): these resources do not produce power on demand. As a result, there is a need for efficient financial incentives that motivate consumers to shift consumption to times when renewable energy is available [2, 3]. Governments around the world are acting to re-engineer their electricity grid into a smart-grid with supporting retail market infrastructure. However, this involves significant risks: the California energy crisis of 2000 and the east-coast blackout of 2003 demonstrated the risks presented by flawed market designs and electricity grids that get unstable.

The Power Trading Agent Competition (Power TAC) is a low-risk platform for modeling and testing retail power market designs and related automation technologies. It simulates a future smart grid environment with renewable energy production, smart metering, autonomous agents acting on behalf of customers and retailers, state-of-the-art customer models, and realistic market designs that include traditional energy exchanges like FERC in conjunction with a modern retail market model that is similar to the Texas energy market, run by ERCOT. The need for power trading agents and for the Power TAC framework is motivated in detail in [1].

In Power TAC, self-interested autonomous broker agents compete with each other to maximize their long-term profits. The simulation advances in 1-hour timeslots, to a total of around 60 simulated days. At each timeslot, brokers try to acquire market share of local producers and consumers (e.g. residential and commercial buildings, wind and solar farms), by publishing *tariffs*, which are contracts for energy selling/procurement under different rates and conditions. A broker can publish a variety of tariffs, including time-of-use tariffs, variable-rate tariffs, and tiered-rate tariffs. Autonomous agents acting on behalf of the customers subscribe to tariffs they find attractive, based on their customers' energy consumption/production profiles. To satisfy customers energy demand, broker agents can bid in a wholesale market to buy electricity from generation companies through sequences of double-auctions. The wholesale market is a day-ahead market: there are 24 double-auctions for each future timeslot, in each of the 24 hours preceding this timeslot. In the electric grid, supply and demand must be balanced at all times. However, balancing becomes harder in the presence of renewable production. Therefore, brokers are financially incentivised to balance supply and demand in their portfolios, by being charged high fees for any imbalance.

This research introduces TacTex'13, the champion agent from the inaugural competition in 2013. TacTex'13 is a complete, fully implemented agent that learns and adapts to the environment in which it operates, by heavily relying on online reinforcement learning and prediction methods. This research formalizes the complex decision-making problem that TacTex'13 faces, and approximates its solution in TacTex'13's constituent components. The success of TacTex'13 is then examined through analysis of competition results.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

2. THE TacTex'13 AGENT

TacTex'13 (referred to here as TacTex) is a utility-maximizing broker agent that operates simultaneously in multiple markets. TacTex's utility measure is the cash amount in its bank account, called its *cash position*. At each timeslot, TacTex executes zero or more actions in both the tariff market and the wholesale market. The executed actions are those that are predicted to maximize its expected long-term utility. In the tariff market, the actions considered by TacTex are consumption-tariff publications, and in the wholesale market the considered actions are bids and asks, to procure and sell energy respectively.

To maximize its utility, TacTex must simultaneously optimize its income and costs and find a long-term profit-maximizing combination of (1) energy-selling prices (2) total energy demand of its customers (controllable by how many customers it agrees to serve), and (3) energy-procurement costs for satisfying this demand. Fully optimizing this combined decision-making problem is intractable; thus, TacTex approximates its solution. Let t be some future timeslot. Let D_t be the energy-demand of TacTex's customers at timeslot t; C_t be TacTex's energy-procurement costs at timeslot t; and p_t be TacTex's energy selling-price at timeslot t. Let $\hat{D}_t, \hat{C}_t, \hat{p}_t$ be the current predictions of D_t, C_t, p_t. Let $u_t(D_t, C_t, p_t) = D_t \times (p_t - C_t)$ be a timeslot-specific utility. Let $\mathcal{A} := A_D \cup A_C \cup p$ be the set of available actions, where here A_D and p are tariffs and price publications, and A_C are wholesale market bids. Let $A_t \subset \mathcal{A}$ be the subset of actions that is taken at timeslot t. TacTex approximates a solution to the following interdependent optimization problems (using '$+i$' to denote 'i timeslots into the future'):

1. Optimize costs given predicted demand:

$$\arg\max_{\{A_{Ct}\}_{t=+1}^{+T}} \sum_{t=+1}^{+T} E[u_t(\hat{D}_t, C_t, \hat{p}_t)] \tag{1}$$

2. Optimize demand and selling prices given predicted costs:

$$\arg\max_{\{A_{Dt}, p_t\}_{t=+1}^{+T}} \sum_{t=+1}^{+T} E[u_t(D_t, \hat{C}_t, p_t)] \tag{2}$$

Thus, instead of optimizing over all possible combinations, we separately optimize demand and costs, each conditioned on the current estimate of the other. Each of the two interdependent optimizations perform local improvement steps, in contrast to a global search for the optimum over all possible combinations, however the gain is a reduction of search complexity from multiplicative to additive. The optimization problems defined by Equations 1, 2 are still intractable. The solutions to these two optimization problems are approximated by TacTex's two constituent components: its wholesale market strategy, and its tariff market strategy.

Wholesale Market Strategy. Using the currently active tariffs, tariff subscriptions, and predicted demand of each customer, TacTex computes the total predicted demand and procures it as cheaply as it can through sequential bidding in the wholesale market, using an online reinforcement learning algorithm. TacTex models the sequential bidding as a Markov Decision Process (MDP) with finite number of states and continuous bidding actions, in a way that allows for efficient reuse of data and computation. The result is a sample-efficient online reinforcement learning algorithm, which minimizes procurement costs.

Tariff Market Strategy. Using its past performance in the wholesale market, TacTex estimates its future costs. TacTex then generates a set of candidate tariffs and predicts the demand resulting from each potential tariff publication, using a supervised learning algorithm that uses the currently published tariffs and their subscriptions as training data, and using energy predictions for each of its customers. If the best tariff is expected to increase utility, TacTex publishes it. Otherwise, no action is taken in the tariff market.

3. Power TAC 2013 FINALS ANALYSIS

The Power TAC 2013 finals included 7 competing brokers developed by research groups from Europe and the USA. The competition included all possible combinations of 2-broker and 4-broker games (21 and 35 games respectively), and 4 7-broker games. In the 2-agent games TacTex won all of its 6 games. In the 4-agent games, TacTex won 15 out of the 16 games it completed successfully (TacTex got disconnected from 4 games due to technical issues with the infrastructure we used). TacTex did not win the 7-agent games despite having the largest volume of customers. At a high level, TacTex's wholesale market strategy and tariff market strategy were responsible for TacTex's success in the finals. The wholesale market strategy maintained low-costs, while the tariff market strategy maximized income given these costs. In the 2-agent games TacTex made 32.4% and 88.2% more profits than the 2nd and 3rd place brokers while maintaining similar levels of income-to-costs ratio (1.97), compared to theirs (2.07 and 2.26). In the wholesale market, TacTex's wholesale market strategy procured 50.5% more energy than the 2nd place broker, at a lower cost-per-kWh. TacTex's tariff market strategy published tariffs that resulted in 39.0% and 113.6% more income compared to 2nd and 3rd place agents. In the 4-agent games, TacTex's tariff market strategy decreased prices in the presence of stronger competition, while maintaining 38.1% and 404.5% more profits than 2nd and 3rd place brokers.

4. CONCLUSIONS

This research introduces TacTex'13, the champion power trading agent from the Power TAC 2013 finals. TacTex'13 implements an approximate solution to a complex decision making problem using a utility optimization approach. Future research directions include investigating the usage and optimization of different types of tariffs and energy balancing methods, as well as their impact on the smart grid.

5. REFERENCES

[1] W. Ketter, J. Collins, and P. Reddy. Power TAC: A competitive economic simulation of the smart grid. *Energy Economics*, 39(0):262 – 270, 2013.

[2] W. Ketter, M. Peters, and J. Collins. Autonomous agents in future energy markets: The 2012 Power Trading Agent Competition. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*. AAAI, 2013.

[3] S. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings. Putting the "smarts" into the smart grid: A grand challenge for artificial intelligence. *Communications of the ACM*, 55(4):86–97, 2012.