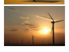
TacTex'13: A Champion Adaptive Power Trading Agent

DANIEL URIELI, PETER STONE
The Univ. of Texas at Austin
Austin, TX 78712 USA
{urieli,pstone}@cs.utexas.edu

Learning Agents Research Group
The University of Texas at Austin

Motivation

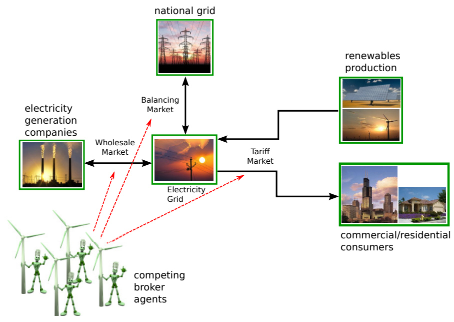
Energy Markets

- Renewable resources do not provide energy on demand
  - Old paradigm: supply follows demand 
  - New paradigm: demand follows supply 
- Needed: new market structures that motivate sustainable behaviors by all participants
- Issues to address: reliability, balancing, peak-demand, energy-efficiency, environmental effects
- However, transition to new market structures can be risky, e.g. California energy crisis 2000-2001

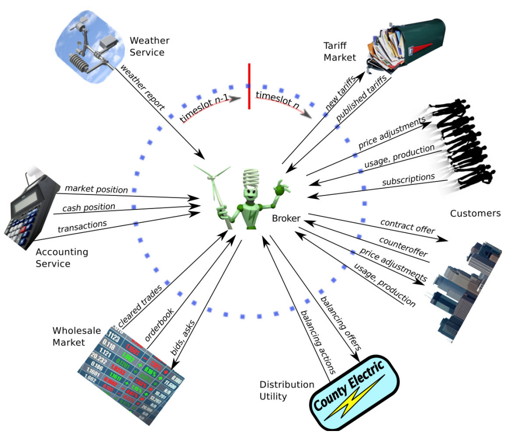
Power TAC

- The Power Trading Agent Competition (Power TAC): open competition using rich smart grid simulation environment
- Focus: retail power markets structure and operation
- Competitors: autonomous broker agents
- Low-risk platform for modeling and testing:
  1. Retail power market designs
  2. Related automation technologies

Power TAC: Game Description

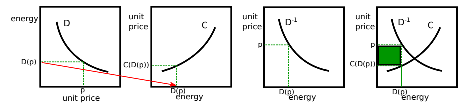


Power TAC: Broker Operation Cycle



TacTex'13 Approach: Demand and Cost Curves

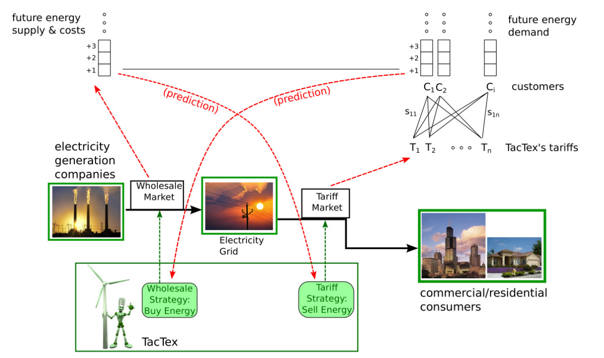
- Goal: maximize profits. How: buy and sell energy
- Define: demand curve D, unit-cost curve C



- Utility $u(p) := D(p) \times p - D(p) \times C(D(p))$
- For a single time step the optimization problem is $\arg\max_{\{A_D, A_C, p\}} u(D, C, p)$, where A_D, A_C, p are the actions affecting the demand-curve, cost-curve and energy selling-price, respectively. The full, sequential optimization problem is $\arg\max_{\{A_t \subset A\}_{t=+1}^{+T}} \sum_{t=+1}^{+T} E[u_t(D_t, C_t, p_t)]$. This problem is generally intractable. We approximate it using a local solution.

Approximate Local Solution

- Optimize costs given predicted demand
- Optimize demand and selling prices given predicted costs

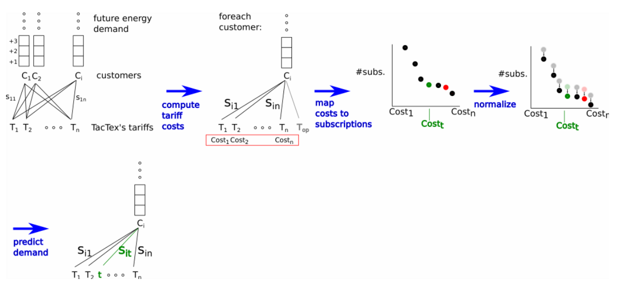


Decision Making in the Tariff Market

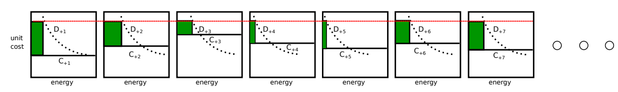
- Primary way to sell energy
- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

- Core computation: "if I publish tariff t, how would it affect my long-term utility?" $\implies$ how would it affect customer subscriptions and demand?



- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence
- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits



Publish tariff if expected to increase utility

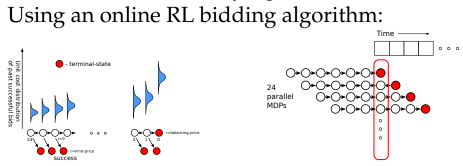
Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:

- Day ahead market $\implies$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\implies$ buy all needed energy

Challenge: what bidding strategy to use?

- For each future timeslot: estimate future demand, then minimize cost for satisfying this demand
- Using an online RL bidding algorithm:



MDP States: $\{0, 1, \ldots, 24, success\}$
MDP Actions: limit-price $\in \mathbb{R}$

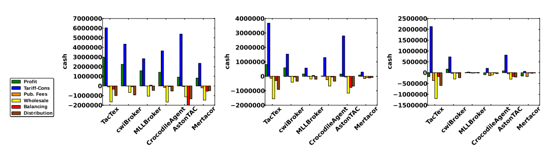
- Transition function is learned from past clearing-prices.
- Strong state generalization allows for quick learning through high reuse of computation and data.

Power TAC 2013 Final Results

- Our agent, TACTEX'13, won the Power TAC 2013 finals:

Broker	7-broker	4-broker	2-broker	Total (not normalized)
TacTex	-705248	13493825	17853189	30641766
cwiBroker	647400	12197772	13476434	26321606
MLLBroker	8533	3305131	9482400	12796064
CrocodileAgent	-361939	1592764	7105236	8336061
AstonTAC	345300	5977354	5484780	11807435
Mertacor	-621040	1279380	4919087	5577427
INAOEBroker02	-76112159	-497131383	-70255037	-643498580

- 2-agent (top-left), 4-agent (top-right), 7-agent games:



Controlled Experiments - Ablation Analysis

- Round-Robin 2-agent tournament between:
  - B: baseline agent
  - U1: adding tariff-market strategy
  - U9_MDP: adding wholesale-market strategy
  - U9_MDP_LWR: adding LWR customer prediction
- Each pair played 200 games with similar conditions

	B	U1	U9_MDP
U9_MDP_LWR	1278.3 (43.2)	708.9 (35.6)	34.2 (23.2)
U9_MDP	966.4 (40.5)	592.6 (22.2)	
U1	547.4 (27.7))		

- 4-agent games using 3 available finalist agents

Broker	Cash
cwiBroker	340.9 (8.4)
Mertacor	-276.2 (40.2)
CrocodileAgent	-287.1 (14.5)
B	**-334.6 (8.0)**

Broker	Cash
cwiBroker	315.4 (9.3)
U1	**135.3 (12.3)**
CrocodileAgent	-372.1 (17.0)
Mertacor	-485.5 (28.1)

Broker	Cash
U9_MDP	**389.9 (13.3)**
cwiBroker	138.3 (8.7)
CrocodileAgent	-333.3 (17.0)
Mertacor	-494.1 (29.6)

Broker	Cash
U9_MDP_LWR	**350.8 (13.3)**
cwiBroker	132.4 (9.0)
CrocodileAgent	-336.9 (17.3)
Mertacor	-566.1 (26.8)

- Tariff and Wholesale strategies improve performance
- LWR customer prediction reduces performance

Summary

- TacTex'13: utility-optimizing broker agent
- Interdependent optimization problems
  - Utility-maximizing tariff strategy
  - Online reinforcement learning bidding algorithm
- Outlook
  - Investigating other tariff, wholesale and balancing strategies
  - Impact on the smart grid and customer behaviors