

Algunas ideas en *machine learning* for the smart grid

Emilio Parrado Hernández
G2PI/DTSC/UC3M

MLG, 20 octubre 2014

Motivación

- ***Smart grid*** es una aplicación que pasa a ser de interés estratégico para el G2PI junto con **neuroimagen e internet as a data source**
- Hay que buscar problemas de *smart grid* que puedan ser resueltos en **menos de 3 años** usando *machine learning*
- Hay que buscar problemas en los que haya datos públicos disponibles.

Contenidos de la charla

- Escenario de aplicación: *smart grid*
- Predicción de consumo
- Desagregación de demanda
- Optimización de consumo
- Datos disponibles
- Discusión

Escenario de aplicación: *smart grid*

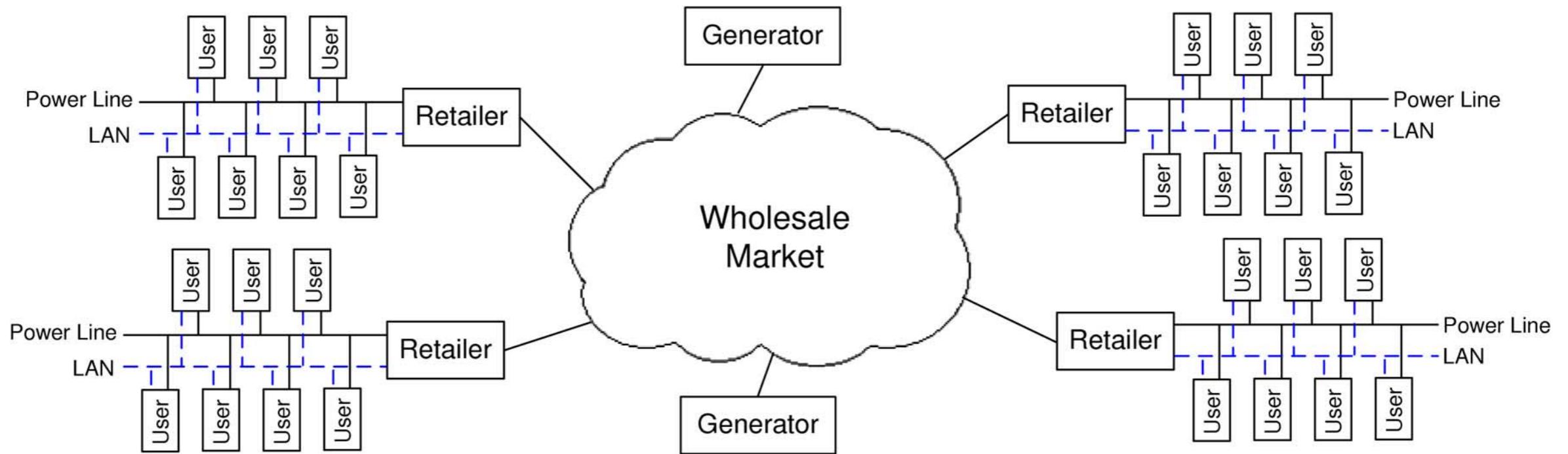
Smart grid

- Red “inteligente” de distribución de energía eléctrica
- Integra: usuarios finales, centros de distribución, generadores de renovables, generadores “tradicionales”, vehículos eléctricos, baterías de almacenamiento.
- Gestión inteligente para:
 - Hacer la demanda plana (p.v. generación)
 - Usar la energía barata (p.v. usuario final)

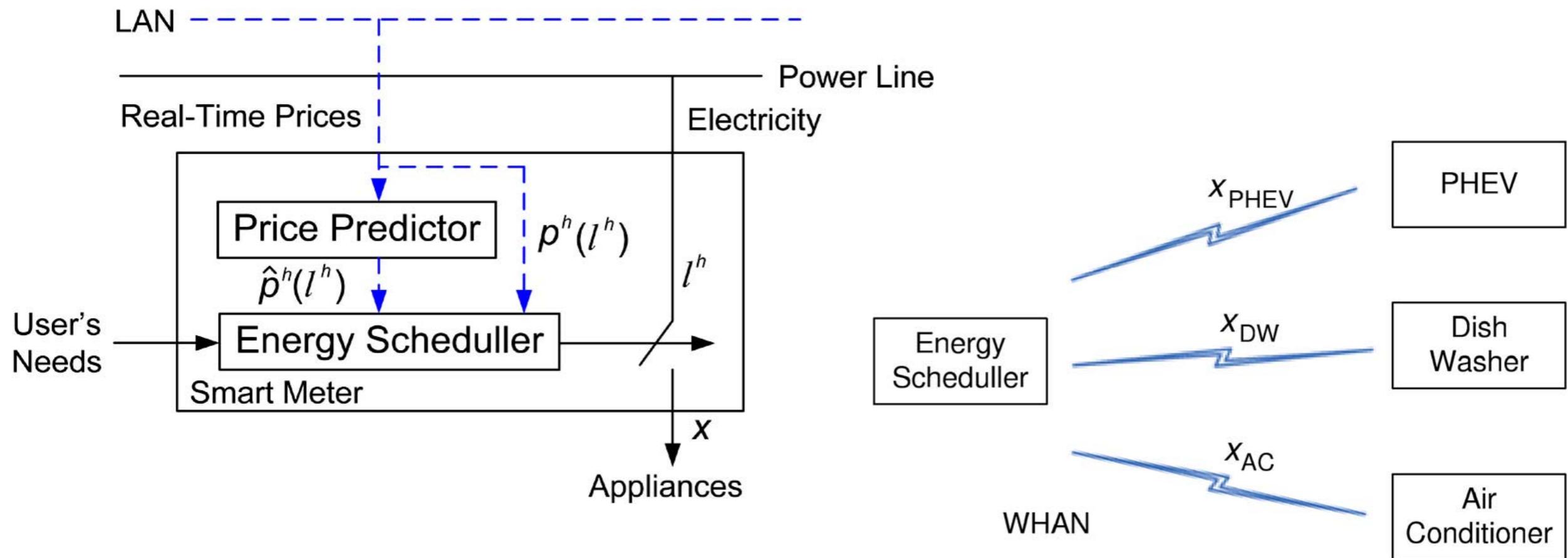
Elementos nuevos

- Esquemas de precio variables con el tiempo para controlar PAR (*peak-to-average-ratio*)
 - *RealTime Pricing*
 - *Inclining Block Rates*. El precio marginal de la energía se incrementa significativamente (40%, 80%) por tramos de consumo
 - *Direct Load Control*: cortar suministro
- *Smart meters* en cada hogar
- Sistemas Domésticos de Gestión de Energía

Mercado de electricidad



Smart meter con planificador



Oportunidades para *Machine Learning*

- Diseño de estrategias de *Real Time Pricing*
- Diseño de algoritmos de Planificación Doméstica
- Componentes para **herramientas de simulación**
 - Modelos de **máquina eléctrica**
 - Modelos de **usuario final**
- Variables exógenas (meteorología, calendario laboral, patrones sociodemográficos, etc).

Predicción de Consumo

Interés en predecir el consumo

- Diferentes granularidades
- Diario: planificación de uso de energía, diseño de señales de precio, previsión de almacenamiento de energía
- Horario o cada 15 minutos: reacción a señales de precio en tiempo variable

Datos a emplear para predecir el consumo

- Consumo de cada hogar (edificio) muestreado cada 15 minutos durante años
- Calendario laboral
- Uso y materiales del edificio
- Meteorología

Artículos

- [Aman 2011] Improving Energy Use Forecast for Campus Micro-grids using Indirect Indicators. IEEE Workshop on Domain Driven Data Mining
- [Mori 2011] Hybrid intelligent method of Relevant Vector Machine and regression tree for probabilistic load forecasting. IEEE intl. Conf. and Exh. on Innovative Smart Grid Tech.
- [Aman 2014]: Holistic Measures for Evaluating Prediction Models in Smart Grids. IEEE tran. on Knowledge and Data Engineering
- [Martinez-Alvarez 2011]: Energy time series forecasting based on pattern sequence similarity. IEEE tran. on Knowledge and Data Engineering

[Aman 2011]

- Regression Trees (implem. matlab):
 - Mezclan variables categóricas con series temporales.
 - permiten analizar relevancia de variables
- Baseline:
 - Diario: Media anual, media del día de la semana (DoW) y media del día del año (DoY)
 - 15 min: media anual, time-of-week, time-of-day

[Aman 2011]

Table I
PREDICTION ERRORS FOR BASELINES AND REGRESSION TREE MODELS FOR DAILY CAMPUS ENERGY USE FORECAST

Model Used	CV-RMSE
Annual Mean	11.32%
Day of Week Mean	14.39%
Day of Year Mean	12.62%
Regression Tree	7.45%

Table III
PREDICTION ERRORS FOR BASELINES AND REGRESSION TREE MODELS FOR 15-MIN CAMPUS ENERGY USE FORECAST

Model Used	CV-RMSE
Annual 15-min Mean	17.37%
Time of Week Mean	16.00%
Time of Year Mean	15.07%
Regression Tree	13.70%

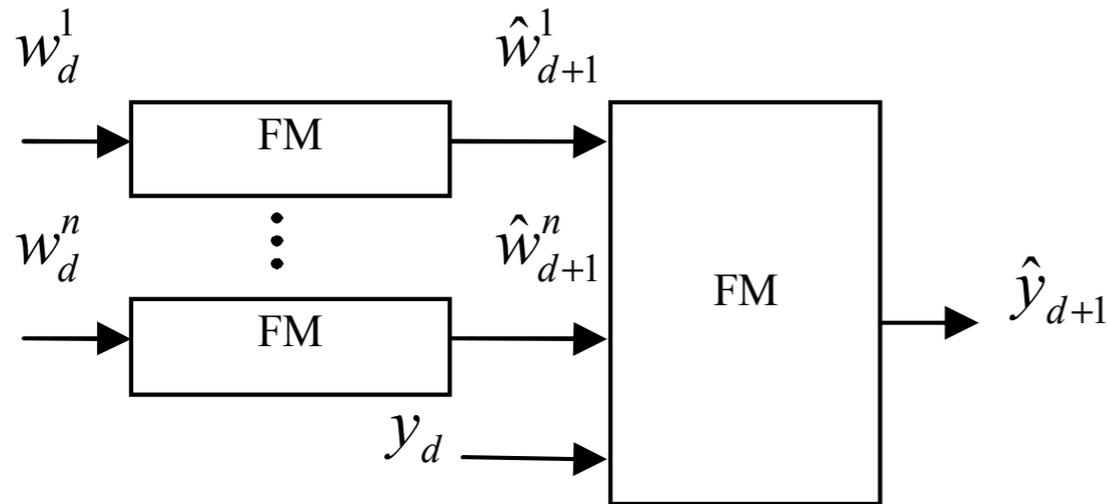
Table II
PREDICTION ERRORS FOR BASELINES AND REGRESSION TREE MODELS FOR DAILY CAMPUS ENERGY USE FORECAST

Weekday	Semester	Temperature	Holiday	CV-RMSE
◆	◆	◆	◆	7.40%
◆		◆	◆	7.60%
◆		◆		7.95%
◆	◆	◆		8.05%
◆	◆		◆	8.37%
◆	◆			8.54%
◆				8.86%
		◆		10.48%
	◆			11.05%
			◆	11.54%

Table V
PREDICTION ERRORS FOR BASELINES AND REGRESSION TREE MODELS FOR DAILY BUILDING ENERGY USE FORECAST

	OHE	DMT	PSB
Annual Mean	20.55	19.77	23.39
Day of Week	26.13	20.09	23.65
Day of Year	24.64	9.62	27.48
Regression Tree (Global)	12.09	11.77	19.32

[Mori 2011]



B. Algorithm

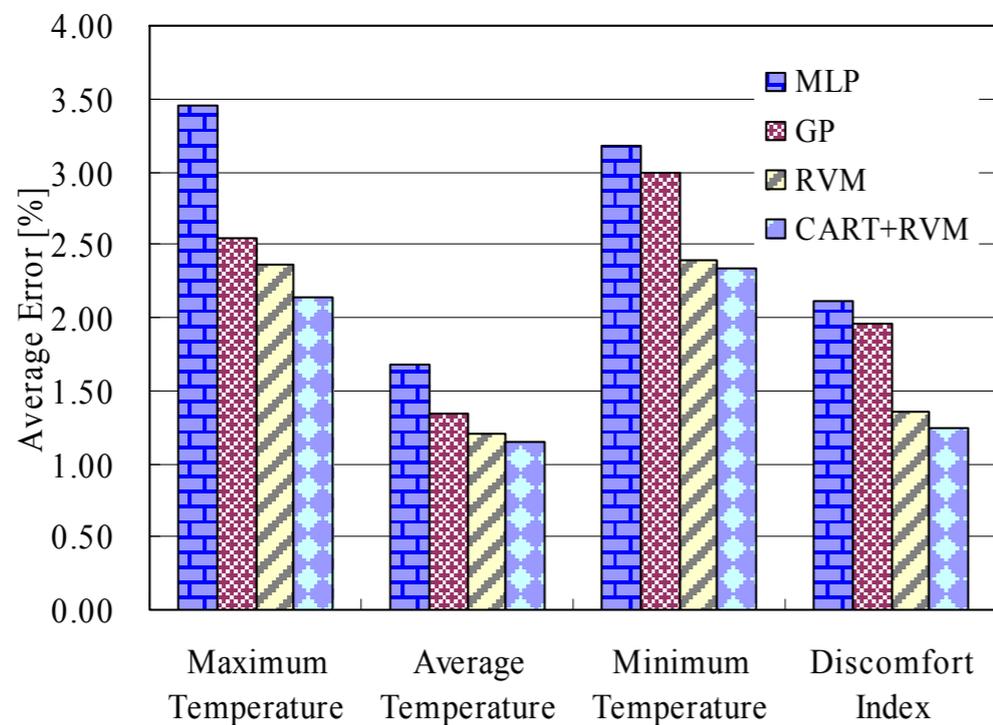
As a whole, the algorithm of the proposed model of CART and RVM may be written as

Step 1: Construct the CART-based regression tree for learning data.

Step 2: Calculate the variable importance of CART and select input variables.

Step 3: Construct the RVM forecasting models at each terminal node of CART with learning data classified by each terminal node.

Step 4: Give unknown data to the proposed model, classify it to the appropriate terminal node, and evaluate the average and the variance of the predicted load with the RVM model.

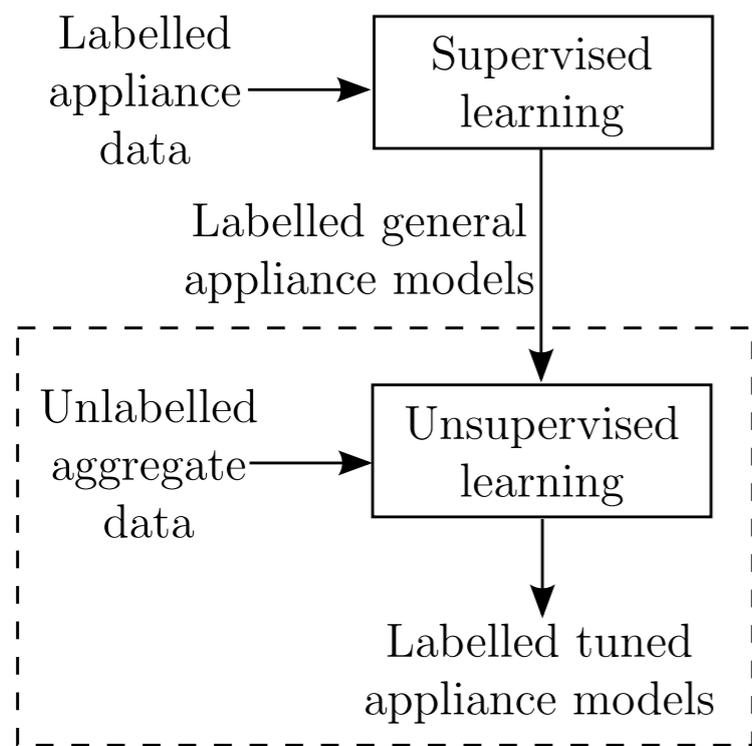


Desagregamiento del consumo de un usuario final

Problema en cuestión

- Desagregar el consumo total de un usuario final en las contribuciones de todos los electrodomésticos.
- Modo **no intrusivo** y con poco etiquetado
- Problema al escalar a hogares de test
- [Parson 2014] An unsupervised training method for non-intrusive appliance load monitoring. Artificial Intelligence 2014.

[Parson 2014]



(a) Our proposed approach

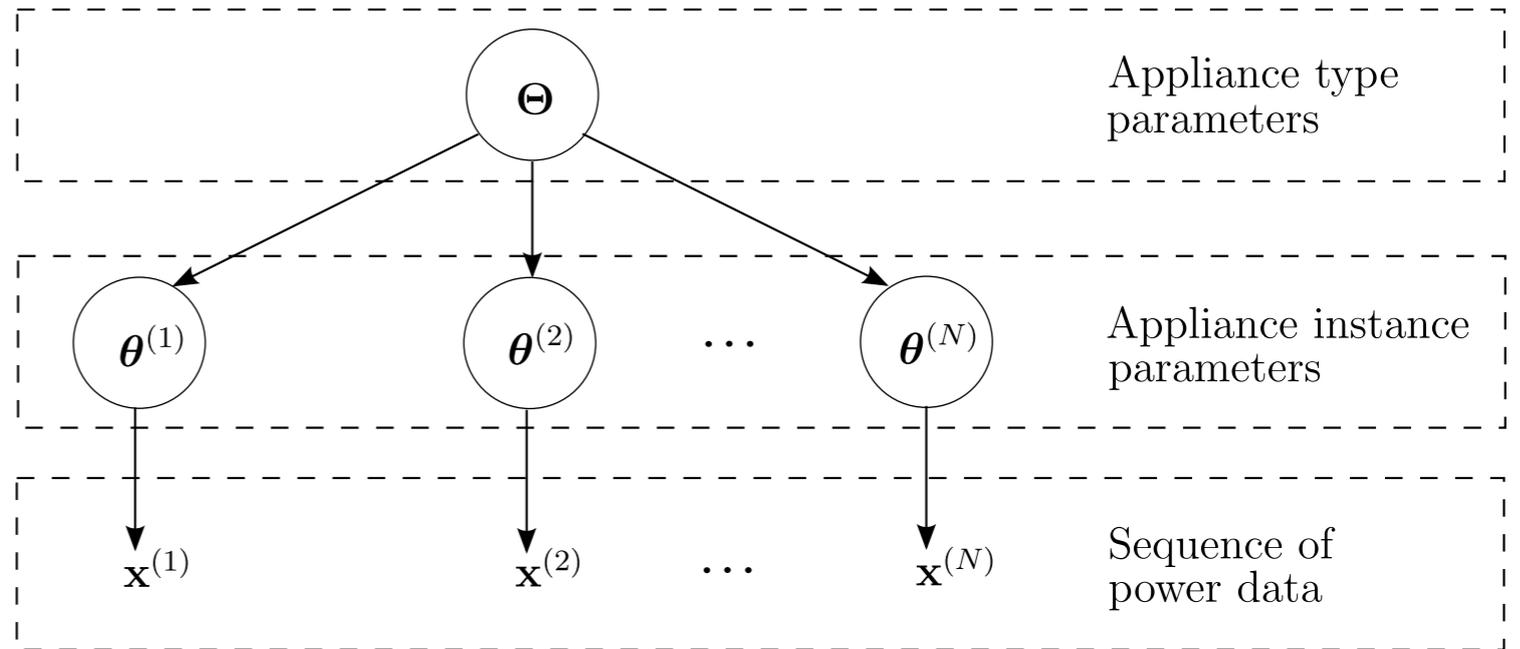
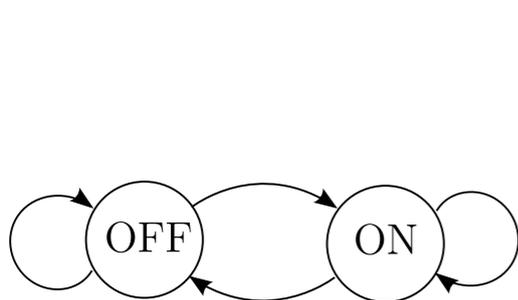
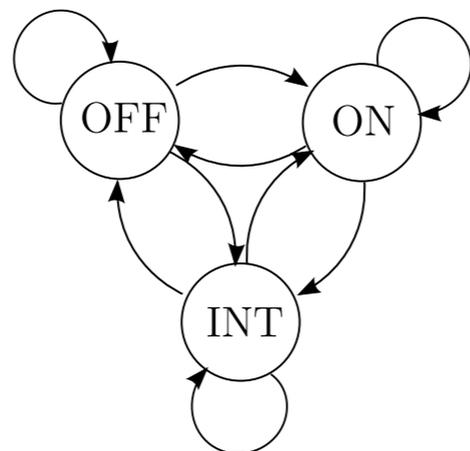


Fig. 2. Hierarchical model of an appliance type Θ .



(a) 2 state model



(b) 3 state model

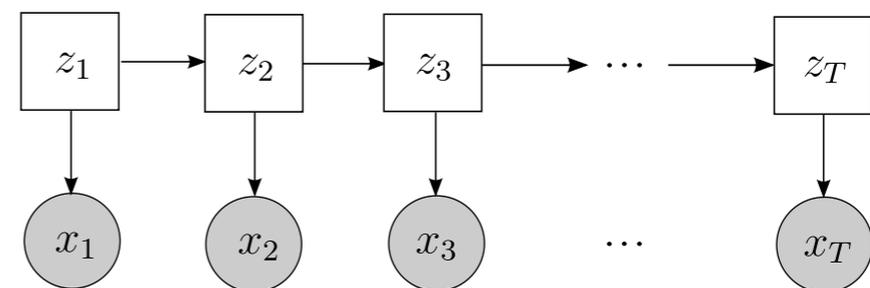


Fig. 4. Appliance state models.

[Parson 2014]

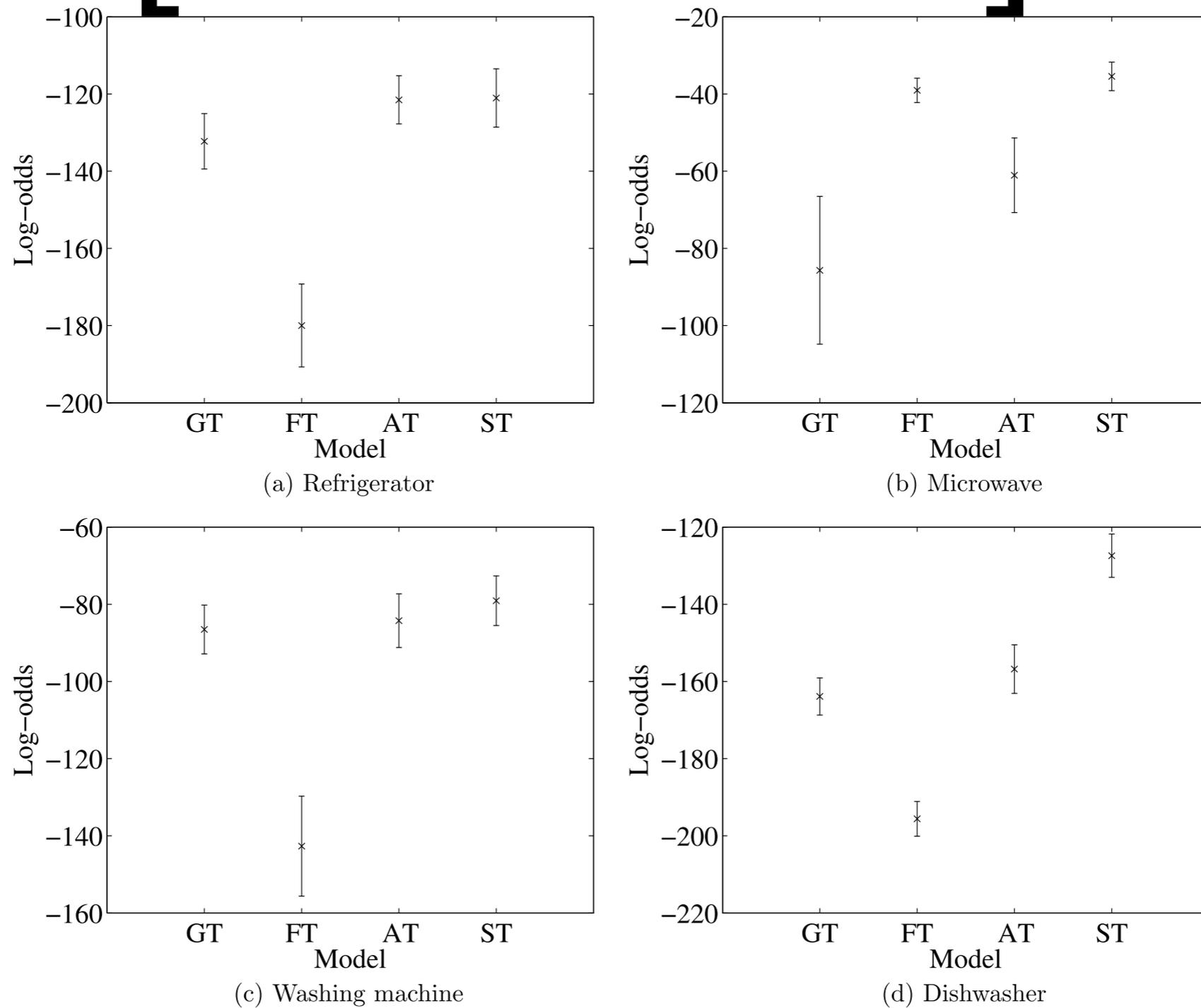


Fig. 10. Mean model log-odds for different training methods. Legend: GT – general model, FT – general model tuned using factorial HMM, AT – general model tuned using extracted signatures, ST – general model tuned using sub-metered data. Error bars represent standard error in the mean.

Planificación de consumo en tiempo real

Planteamiento del problema

- Respuesta a los esquemas de Respuesta a la Demanda y Tarificación en Tiempo Real desde los hogares.
- Aparecen *schedulers*: sistemas domésticos de gestión de energía (SDGE) conectados al *smart meter* y a las máquinas eléctricas.
- SDGE: Problema de optimización con restricciones: minimizar coste satisfaciendo necesidades
- Necesidades: Algoritmo de SDGE, y estimación de datos de entrada

Artículos relacionados

- [Mohsenian 2010] Optimal Residential load control with price prediction in real-time electricity pricing environments. IEEE tran. on Smart Grid 1(2), 2010
- [Mohsenian 2010] Autonomous demand-side management based of game-theoretic energy consumption scheduling for the future smart grid. IEEE tran. on Smart Grid 1(3), 2010
- [Chen 2012] Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. IEEE tran. on Smart Grid 3(4), 2012
- [Wu 2014] Real-time Scheduling of Residential Appliances via conditional risk-at-value. IEEE tran. on Smart Grid 5(3), 2014

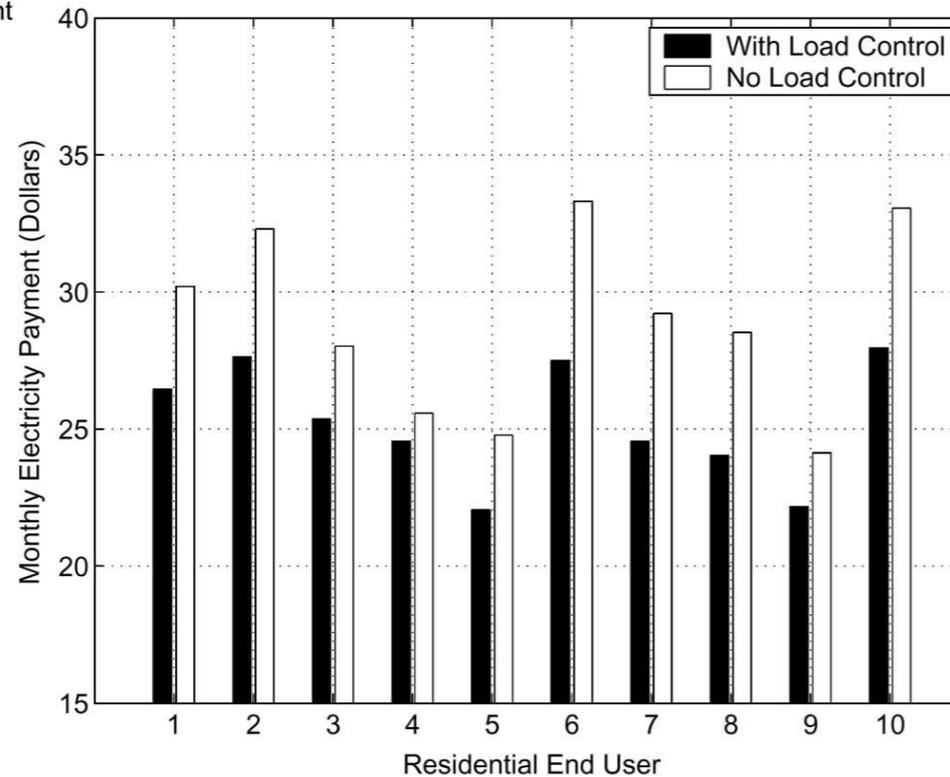
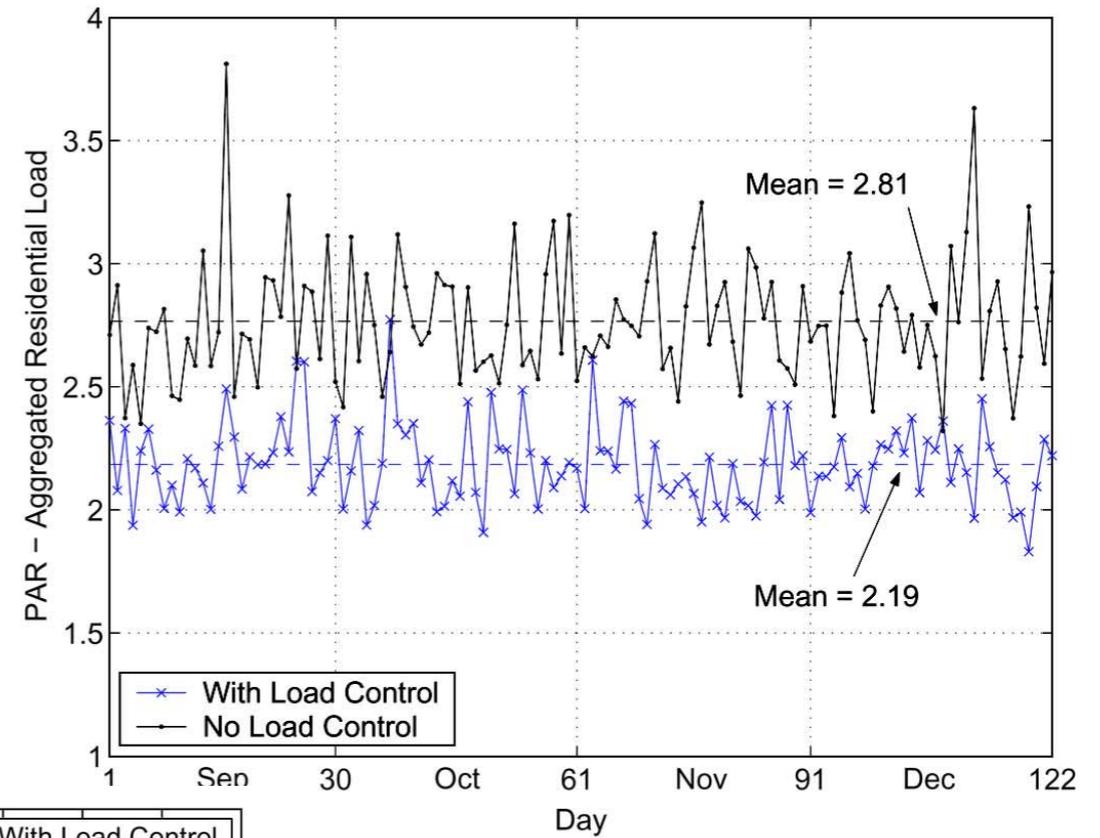
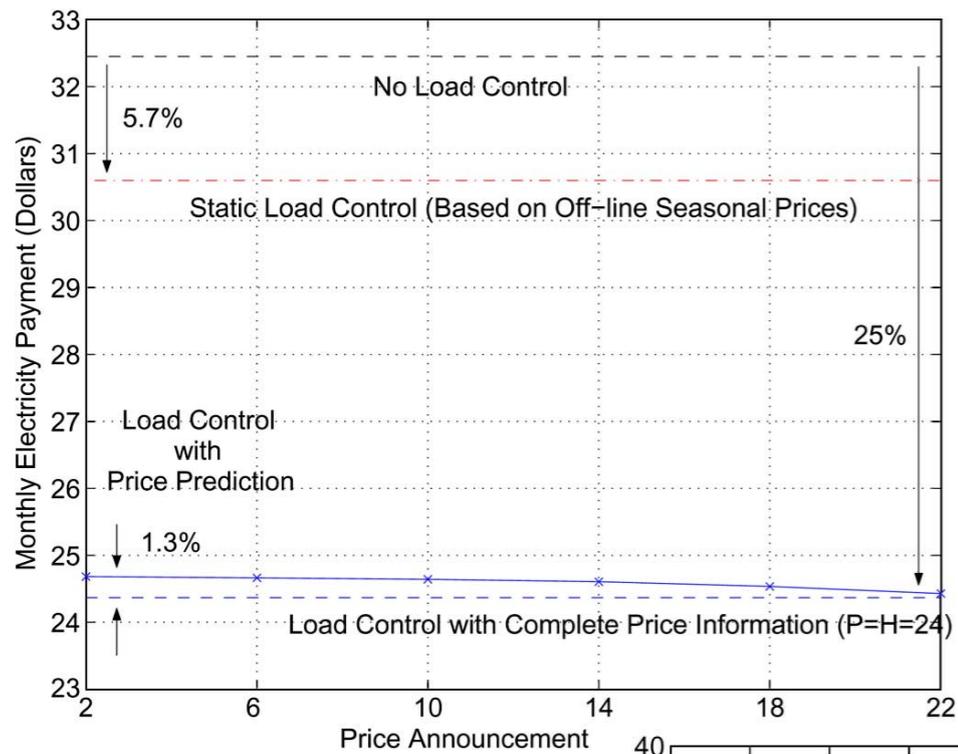
[Mohsenian 2010]

- Respuesta a tarificación IBR o RTP:
- Optimización:
 - Variables: vector por cada electrodoméstico cuyas componentes dicen cuánto consume en cada instante de tiempo
 - Restricciones: suma de consumo total por hora, suma del consumo de cada elect., tiempo de uso de cada electrodom.
 - Funcional: Precio a pagar en todo el día usando tarificación variable y tiempo de espera para encender cada electrodoméstico

[Mohsenian 2010]

- **Predicción de la señal de precio:** para escenarios donde el precio en cada hora se anuncia 5-6 horas antes (no un día antes)
- Usar info a priori: modelo lineal sencillo con datos de ayer, antes de ayer y mismo día la semana pasada para reducir coste computacional

[Mohsenian 2010]



[Mohsenian 2010b]

- Mismo escenario pero ahora los usuarios están conectados entre sí, y pueden aplicar algoritmos colaborativos/distribuidos.
- Optimización simultánea usando teoría de juegos
- Hay un equilibrio de Nash en el que ningún usuario se beneficia de salirse de su vector de consumo óptimo.

[Mohsenian 2010b]

Algorithm 1: Executed by each user $n \in \mathcal{N}$.

- 1: Randomly initialize \mathbf{l}_n and \mathbf{l}_{-n} .
 - 2: **Repeat**
 - 3: **At** random time instances **Do**
 - 4: Solve local problem (29) using IPM [27].
 - 5: **If** \mathbf{x}_n changes compared to current schedule **Then**
 - 6: Update \mathbf{x}_n according to the new solution.
 - 7: Broadcast a control message to announce \mathbf{l}_n to the other ECS units across the system.
 - 8: **End**
 - 9: **End**
 - 10: **If** a control message is received **Then**
 - 11: Update \mathbf{l}_{-n} accordingly.
 - 12: **End**
 - 13: **Until** no ECS unit announces any new schedule.
-

[Mohsenian 2010b]

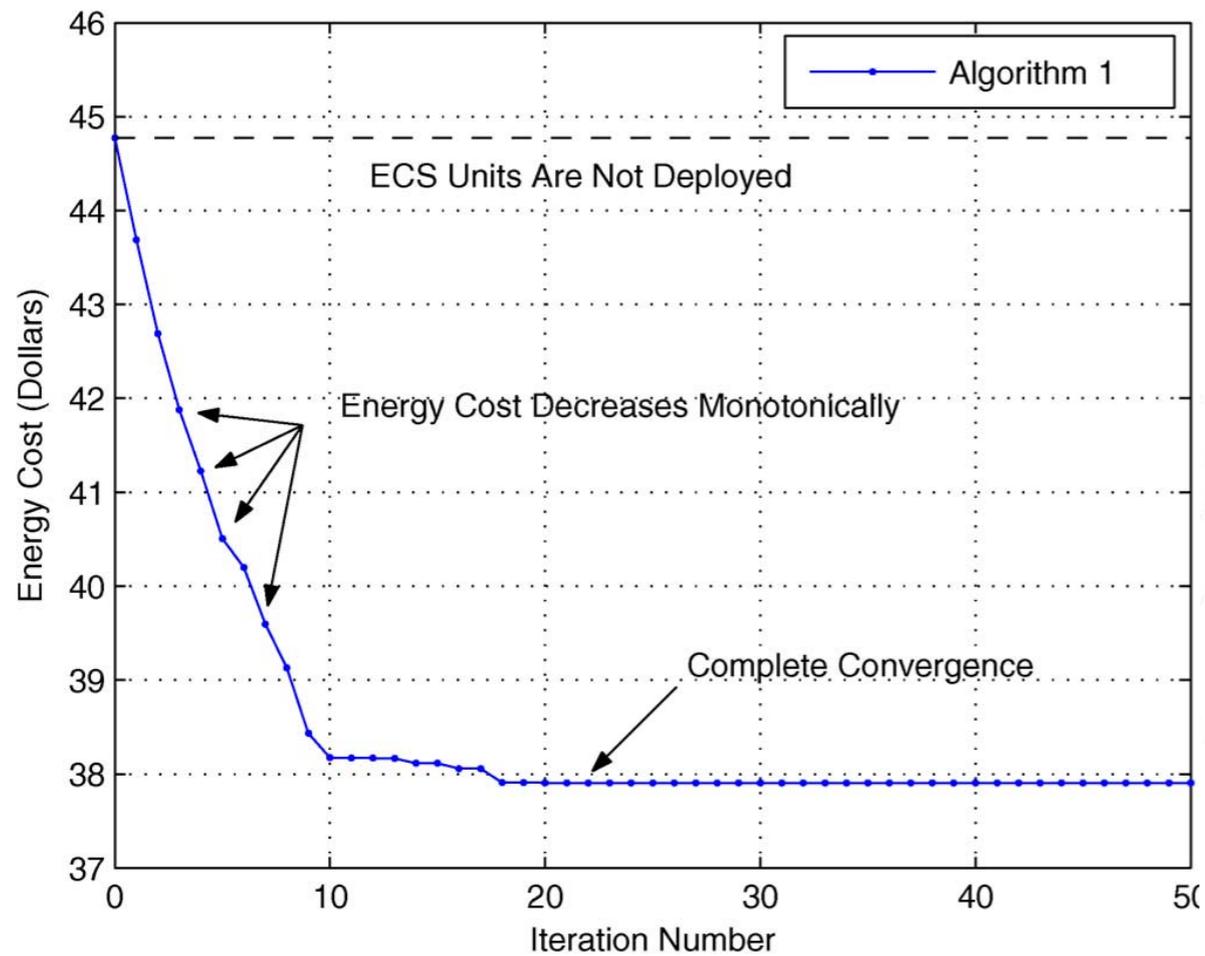
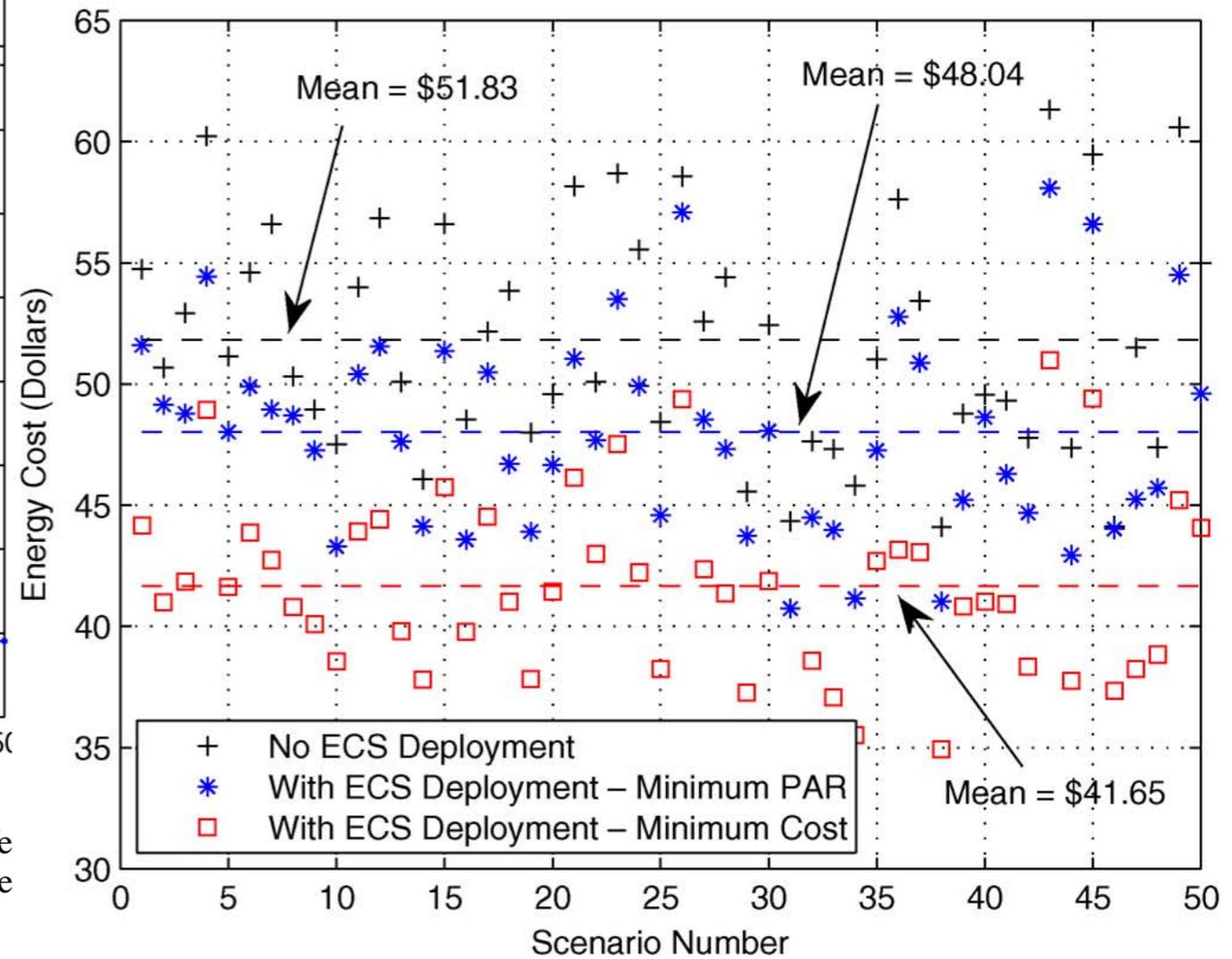


Fig. 8. Trend of resulting energy cost along the iterations of Algorithm 1. We can see that the proposed distributed algorithm converges quickly. Steady state is reached after only 22 iterations when the energy cost is minimized.



[Chen2012]

- Scheduling cada 5 minutos.
- Equilibrio entre coste y riesgo en incertidumbre de precios futuros
- Simulaciones Monte Carlo cada 5 minutos para ver la evolución de precios en el resto del día y actualizar el scheduling.
- Modelos de señales de precios: Gaussianas
- Optimización con riesgos y sin riesgos

[Wu 2014]

- Idea de optimizar el equilibrio entre precio y riesgos futuros (incertidumbre en el precio)
- Usa un controlador basado en lógica borrosa para el optimizador.

Bases de Datos

[Aman 2011]

- Datos de smart meters en 170 edificios del campus de USC durante 3 años. Muestreo cada 15 minutos
- Categorías de cada edificio: Academic, Residential, Other
- Datos de meteorología: clima, humedad, cada hora
- Superficie neta y bruta, año de construcción para caracterizar el edificio
- Patrones de uso de cada edificio según calendario académico

[Parson 2014]

- Tracebase: 1270 trazas de 122 dispositivos.
- REDD: Reference Disaggregation Data set. Registros de consumo de energía individuales para todos los electrodomésticos (268 monitores) de 10 hogares durante 119 días (combinados en todas los hogares), supone más de 1 Terabyte de datos crudos.

[Mohsenian 2010,b]

- Tarifas RTP de Illinois Power Company de enero 2007 a diciembre 2009
- Generan casas con entre 10 y 25 máquinas eléctricas, incluyen consumo fijo, horno, bombillas (10), calentador, lavadora, coche eléctrico, etc

[Wu 2014]

- [25] Solar Energy Of Illinois [Online]. Available: http://67.162.55.142/_a332427bda2b71d210ff16a738c38940_/?s=reportshow&i=day-power
- [26] Department of Atmospheric Sciences, University of Illinois [Online]. Available: <http://www.atmos.illinois.edu/weather/daily/index.html>
- [27] Annex 42 [Online]. Available: <http://www.iea-annex54.org/annex42/index.html>
- [28] Ameren Illinois [Online]. Available: <http://www.ameren.com/sites/aiu/ElectricChoice/Pages/ResRealTimePricing.aspx>
- [29] RELOAD Database Documentation and Evaluation and Use in NEMS [Online]. Available: <http://www.onlocationinc.com/LoadShapes-Reload2001.pdf>

Discusión

Ideas para seguir

- Predicción de consumo: métodos no muy sofisticados, análisis más o menos sencillo y publicado en IEEE tran. on Knowledge and data Engineering.
- Problemas teóricos en el desagregado de consumo: etiquetado difícil, inferencia bayesiana
- Necesidad de optimización y simulación de escenarios distribuidos para el *scheduling*