

Técnicas de Predicción para la Eficiencia Energética en Smart Cities

Fernando J. Isorna¹, Alicia Troncoso¹

¹ Departamento de Ingeniería Informática, Universidad Pablo de Olavide
fjisoret@alumno.upo.es, atrolor@upo.es

1 Resumen

Este proyecto de tesis se enmarca en el contexto de la imperiosa necesidad de reducir el consumo energético de nuestras ciudades. La tesis parte de la base de que, en la actualidad, resulta relativamente fácil y económico monitorizar el consumo energético de los edificios. Con esta monitorización se obtienen ingentes cantidades de datos que nos aportarán una valiosa información sobre los perfiles de consumo, entre otros. Así, con esta tesis se pretende extraer conocimiento a partir de esos datos, de forma que se pueda predecir el consumo futuro de los edificios en función de algunas variables como por ejemplo: la meteorología, perfil de uso, etc. Además, se podrán comparar los consumos entre edificios parecidos, de forma que se puedan detectar anomalías que permitan corregir hábitos de consumo energético inadecuados. Por otro lado, se persigue poder proponer soluciones que lleven a un uso más eficiente de la energía y monitorizar adecuadamente el resultado de las soluciones que se implanten.

1.1 Introducción

Uno de los retos en el contexto de Smart Cities es la mejora de la eficiencia energética a través de la gestión inteligente de edificios, del alumbrado público, del abastecimiento de agua, etc. Hoy en día debido al uso de sensores y contadores inteligentes, tanto en el ámbito doméstico como en los edificios públicos, se puede extraer de una forma fácil los datos de consumo de energía con el objeto de descubrir conocimiento acerca de los hábitos de consumo de los usuarios para así generar predicciones que ayuden a mejorar la eficiencia energética.

La predicción de eventos futuros siempre ha fascinado al ser humano y se puede decir que las técnicas de predicción han sido objeto de estudio desde que éste existe. En la actualidad, son numerosas las actividades humanas y fenómenos naturales que producen grandes volúmenes de datos que evolucionan con el tiempo. La aparición primero de la estadística y posteriormente de la informática ha hecho factible el estudio de estos datos temporales. A los valores de una variable en distintos momentos de tiempo se le conoce con el nombre de Serie Temporal. Como es sabido, hoy en día la capacidad para generar datos es abrumadora debido a los grandes avances tecnológicos, pero no así la capacidad para interpretarlos y extraer información y conocimiento útil de los mismos. Por tanto, las Series Temporales son un gran reto dentro de esta área de investigación.

Los datos temporales son un t3pico de investigaci3n en diferentes disciplinas desde hace d3cadas. Sin embargo, las soluciones tradicionales distan mucho de resolver los retos actuales. Los datos temporales consistentes en algunos miles o decenas de miles de valores de una variable indexada cada hora es historia. En la actualidad no s3lo debemos trabajar con series de datos multidimensionales con decenas o centenares de atributos interrelacionados entre s3, sino que su escala temporal puede ser el segundo. Por tanto, es preciso desarrollar nuevas t3cnicas avanzadas que ayuden a descubrir las principales caracter3sticas de una serie temporal con el objetivo principal de predecir su comportamiento en un futuro cercano.

Por ello, se plantea como un gran objetivo de esta tesis la posibilidad de desarrollar modelos avanzados de predicci3n a partir de patrones de consumo con el objeto de la mejora de la eficiencia energ3tica en edificios p3blicos en el contexto de Smart Cities. Es de gran inter3s hoy en d3a poder obtener predicci3nes de consumo desagregada, por hogares en el caso de las empresas el3ctricas o por edificios p3blicos en el caso de instituciones p3blicas, lo m3s aproximadas posibles para que el ajuste de la producci3n al consumo en tiempo real permita el ahorro energ3tico.

1.2 Hip3tesis

Se puede decir que el intento de predecir fen3menos del futuro con m3s o menos acierto, se ha llevado a cabo desde siempre. Obtener una predicci3n aproximada es crucial en actividades econ3micas ya que errores en la predicci3n conllevan una gran p3rdida econ3mica.

Tradicionalmente, las t3cnicas de predicci3n han estado basadas en modelos estad3sticos tales como modelos AR, MA, ARMA, ARIMA o GARCH [1], o en t3cnicas de Inteligencia Artificial como las redes neuronales [2], m3quinas de vectores soporte [3], modelos basados en reglas [4] o modelos h3bridos que intentan destacar las ventajas de cada uno de los m3todos [5]. Sin embargo, estas t3cnicas se muestran insuficientes sobre todo cuando se trata de resolver problemas complejos con datos del mundo real y de grandes dimensiones. En los 3ltimos a3os, se han aplicado t3cnicas de Miner3a de Datos [6, 7, 8, 9] tradicionalmente usadas para problemas de clasificaci3n o clustering, a la predicci3n de datos temporales. De esta manera, dichas t3cnicas han sido adaptadas a problemas de regresi3n para resolver problemas concretos tales como la predicci3n de la demanda de energ3a y la predicci3n de los precios de la energ3a en mercados el3ctricos de diferentes pa3ses, mostrando resultados muy competitivos. Por ejemplo en [9] se aplic3 un m3todo de regresi3n basado en vecinos cercanos, donde previamente se establec3a el tama3o 3ptimo de una ventana temporal formado por valores pasados y el n3mero 3ptimo de vecinos.

En los 3ltimos a3os, el uso de los patrones de comportamiento como paso previo a la predicci3n ha dado muy buenos resultados. En [10] se presenta el algoritmo LBF, que a partir de los patrones de comportamiento de la serie temporal es capaz de mejorar de forma importante los errores de predicci3n de la literatura. Este algoritmo es mejorado en [11] en el que se determina el n3mero de patrones de comportamiento a extraer a trav3s de un sistema de voto mayoritario y el tama3o de la ventana temporal se calcula a trav3s de validaci3n cruzada.

Por otro lado, para obtener patrones de comportamiento a partir de series temporales se necesita implícitamente determinar cuando dos series temporales son iguales o al menos “cercanas” mediante el establecimiento de una distancia. Actualmente, este problema es un reto abierto, sobre todo si el número de muestras es grande. La aproximación MUSS [3] es capaz de calcular distancias entre series temporales con un coste computacional muy inferior a medidas clásicas como DTW [12].

Aunque hay muchos trabajos publicados recientemente acerca de técnicas de predicción aplicadas al consumo energético, existe un vacío en lo que respecta a predicción en el contexto de Smart Cities. En este contexto, es importante obtener predicciones desagregadas para cada edificio o para cada hogar principalmente a partir de datos obtenidos de contadores inteligentes o redes de sensores. Por tanto, en esta tesis se pretende cubrir este vacío con el desarrollo de modelos de predicción a partir de los datos de consumo energético procedentes de un conjunto de edificios públicos. Estos modelos están basados en técnicas avanzadas de la Inteligencia Artificial y tienen en cuenta los patrones de consumo de los edificios con el objeto de mejorar la eficiencia energética.

1.3 Objetivos

El principal objetivo de esta tesis es el estudio de modelos de predicción con el objeto de aplicarlos a la mejora de la eficiencia energética en un conjunto de edificios públicos. Una relación de los objetivos se presenta a continuación.

- a. **Estudio de modelos basados en técnicas de regresión para el análisis de datos temporales:** El primer objetivo consiste en un estudio exhaustivo de modelos existentes en la literatura para su posible adaptación a la predicción de consumo de energía de un edificio. Entre estos modelos se estudiarán en especial las técnicas de regresión que conllevan un paso previo de reconocimiento de patrones, debido a los resultados competitivos obtenidos de la aplicación de dichos métodos a problemas reales.
- b. **Perfil de consumo energético de un edificio:** Se analizarán las variables que caracterizan el tipo de actividad de un edificio y la influencia de variables meteorológicas, como pueden ser la temperatura y la radiación solar, en el consumo energético del edificio con el objetivo final de obtener una caracterización individual de cada edificio. También se estudiará en qué forma se puede representar dicha información para que identifique el perfil de consumo de cada edificio. Por otro lado, se podrán comparar edificios según su perfil de consumo energético atendiendo al concepto de distancia entre dos series temporales.
- c. **Patrones de consumo energético:** Una vez analizada la información relevante y obtenidos los perfiles de consumo energético de cada edificio, se aplicarán diferentes técnicas de clustering para obtener una agrupación de edificios según dichos perfiles. De esta forma, obtendremos patrones de consumo de los edificios dependiendo de los días, tipo de actividad de los edificios, etc. Es importante obtener estos patrones pues nos indicarán en qué grupo de edificios es más importante actuar para mejorar la eficiencia energética.

- d. Predicción de consumo energético de cada edificio:** Una vez obtenidos los patrones de consumo de los diferentes edificios, se pretende desarrollar métodos de predicción que exploten esta información para obtener predicciones más aproximadas. Además, se estudiarán diferentes horizontes de predicción, en particular, se harán predicciones de consumo en el muy corto plazo, es decir horas, en el corto plazo de un día y en el largo plazo de una semana.

2 Metodología y Plan de Trabajo

Para asegurar la consecución de los objetivos propuestos, se desarrolla un plan de trabajo descompuesto en las tareas especificadas a continuación:

Revisión del estado del arte: Se realizará un estudio de los distintos tipos de técnicas usadas en la literatura para la reducción de dimensionalidad, agrupamiento y predicción. En particular, se estudiarán los trabajos publicados en los últimos años en las revistas de impacto y en los congresos internacionales más prestigiosos dentro de la Inteligencia Artificial.

Obtención de datos: En esta tarea se acometerá todo el proceso de obtención de datos y su preprocesado. Los datos de consumo energético de cada edificio serán mediciones cada 15 minutos y se obtendrá una muestra de 105120 mediciones correspondientes a los años 2011, 2012 y 2013. Las técnicas de preprocesado consistirán principalmente en la sustitución de los valores perdidos.

Selección de variables: En una primera fase se acometerá el desarrollo de modelos teniendo en cuenta nada más que el consumo energético de cada edificio. Sin embargo, en una fase posterior se definirá una variable que codifique el tipo de actividad del edificio y se estudiará la influencia de la temperatura y la radiación solar en cada edificio para incluir dichas variables en aquellos edificios en los que efectivamente se vea una dependencia. Por último, se aplicarán técnicas de selección de atributos para obtener qué mediciones del pasado son importantes a la hora de obtener una predicción futura.

Reconocimiento de patrones: Una vez definidas las variables, se utilizarán todos los datos de los distintos edificios para obtener patrones de comportamiento energético según el tipo de actividad del edificio y según los días y las horas del día. Estos patrones definirán grupos de edificios con un nivel de eficiencia energética similar y las medidas que se puedan tomar para el ahorro de energía en un edificio serán aplicables a todos los edificios que pertenezcan al mismo grupo.

Técnicas de predicción y resultados experimentales: En esta tarea se aplicarán técnicas de predicción que tengan en cuenta los patrones de consumo energético de los diferentes grupos de edificios bajo estudio. Se desarrollará un modelo de predicción adecuado para cada grupo de edificios que se haya obtenido en la tarea anterior.

Los resultados que se obtengan demostrarán si el uso de variables adicionales al consumo energético mejora el error de la predicción y si obtener diferentes modelos de predicción según el cluster al que pertenezca el edificio es una buena estrategia.

3 Relevancia

Debido a los graves problemas medioambientales que se padecen por la dependencia existente en todo el mundo de la producción de energía, es imperativo reducir al máximo su consumo. Como en otros ámbitos, es fundamental conocer y comprender los patrones de comportamiento, de forma que se pueda plantear una estrategia eficaz para la eficiencia energética en edificios públicos de ciudades.

4 Bibliografía

1. Jan G. De Gooijer, Rob J. Hyndman. 25 years of time series forecasting. *International Journal of Forecasting*, Vol. 22, no. 3, pp. 443-473, 2006
2. E. Gómez-Ramírez, K. Najim, E. Ikonen. Forecasting time series with a new architecture for polynomial artificial neural network. *Applied Soft Computing*, Vol. 7, no. 4, pp. 1209-1216, 2007
3. Alicia Troncoso, Marta Arias, José C. Riquelme. A multi-scale smoothing kernel for measuring time-series similarity. *Neurocomputing*, in press.
4. Martínez-Ballesteros M., Troncoso A., Martínez-Álvarez, Riquelme J. C. Mining quantitative association rules based on evolutionary computation and its application to atmospheric pollution. *Integrated Computer-Aided Engineering*, Vol. 17(3), pp. 227-242, 2010.
5. O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L.J. Herrera, A. Guillen, L. Marquez, M. Pasadas. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems*, Vol. 159, no. 7, pp. 821-845, 2008.
6. A. Ukkonen, M. Arias. Example-dependent Basis Vector Selection for Kernel-based Classifiers. *ECML'2010 European Conference of Machine Learning*, 2010.
7. H. Becker, M. Arias. Real-time Ranking with Concept Drift Using Expert Advice. *13th ACM SIGKDD Internat. Conf. on Knowledge Discovery and Data Mining*, 2007.
8. M. Arias et al. Predicting Electricity Distribution Feeder Failures Using Machine Learning Susceptibility Analysis. *AAI'06 8th Conference on Innovative Applications of Artificial Intelligence*, 2006.
9. A. Troncoso, J. M. Riquelme Santos, A. Gómez Expósito, J. L. Martínez Ramos, J. C. Riquelme. Electricity Market Price Forecasting Based on Weighted Nearest Neighbors Techniques. *IEEE Transactions on Power Systems*, Vol. 22, No. 3, pp. 1294-1301, 2007
10. F. Martínez-Alvarez, A. Troncoso, J. C. Riquelme, J. S. Aguilar-Ruiz. LBF: A Labeled-Based Forecasting Algorithm and its Application to Electricity Price Time Series. *ICDM'08 IEEE International Conference on Data Mining*.
11. F. Martínez-Álvarez, A. Troncoso, J. C. Riquelme, J. S. Aguilar-Ruiz. *Energy*

time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge Data Engineering*, Vol. 23, no. 8, pp. 1230-1243, 2011

12. Shankar Vembu, Alexander Vergara, Mehmet K. Muezzinoglu, and Ramón Huerta. On time series features and kernels for machine olfaction. *Sensors and Actuators B: Chemical*, 174:535–546, 2012.