



MODELADO INTELIGENTE DE PATRONES DE MICROCONSUMOS DOMÉSTICOS DE AGUA CON UN ALGORITMO CART

Iago García Suárez¹, Sarai Díaz García², Javier González Pérez³, José A. Olivas⁴

^{1,4}Escuela Superior de Informática Universidad de Castilla-La Mancha, Paseo universidad 4, 13071, Ciudad Real, España

^{2,3}ETSI Caminos, Canales y Puertos Universidad de Castilla-La Mancha, Avda. Camilo José Cela s/n, 13071, Ciudad Real, España

²*Sarai.Diaz@uclm.es*

RESUMEN

El consumo condiciona el movimiento del agua a través de las redes de abastecimiento de agua potable. El despliegue de contadores inteligentes o *smart meters* supone una oportunidad para conocer el consumo por vivienda y caracterizar los microconsumos o consumos por parte de cada aparato o electrodoméstico (uso final) que consume agua dentro de la vivienda. Los estudios orientados a clasificar microconsumos suelen trabajar con sensores de alta resolución, que miden los pulsos de consumo por debajo del litro y cada poco segundos. Este trabajo presenta una metodología que permite identificar y clasificar los microconsumos domésticos de agua potable a partir de señales con resolución por litro y por minuto, más habituales en redes reales. El método consiste en extraer eventos individuales que luego son clasificados en usos finales mediante un algoritmo de aprendizaje basado en árboles de decisión (CART). Los resultados de este primer análisis con datos simulados alcanzan una eficacia de la clasificación del 93% de exhaustividad a nivel global, que aumenta en los eventos continuos asociados a un gran volumen de agua.

Palabras clave

Modelización demandas, usos finales, inteligencia artificial, árbol de decisión

1. INTRODUCCIÓN

La demanda de agua por parte de la población determina el flujo a través de las redes de abastecimiento de agua potable. Conocer qué condiciona y cómo se produce el consumo de agua es por tanto determinante para simular el funcionamiento hidráulico de estas infraestructuras críticas. Son muchos los trabajos que han abordado la caracterización, modelización e incluso predicción de demandas en redes de abastecimiento a diferentes escalas espaciales y temporales en las últimas décadas (e.g., [1, 2, 3]). Una de las aplicaciones que más ha evolucionado en los últimos años gracias al despliegue de contadores inteligentes

(*smart meters*) tiene que ver con la caracterización de usos finales o microconsumos, término referido a aquellos consumos que se producen en cada aparato o electrodoméstico que consume agua dentro de una vivienda (por ejemplo, grifos, duchas o lavadoras, entre otros).

Los estudios para la caracterización de microconsumos pueden basarse en mediciones intrusivas (sensores en los distintos usos finales) o mediciones no-intrusivas (un único equipo de medida a la entrada de la vivienda). Las mediciones intrusivas proporcionan información detallada sobre cada uso final, pero no se emplean con frecuencia debido a su elevado coste y la habitual reticencia de los usuarios. No disponer de medidas dentro de la vivienda supone tener que desarrollar técnicas que permitan desagregar el consumo de la vivienda en sus correspondientes usos finales. Estas técnicas pueden ser totalmente automáticas o tener una componente manual (basada en juicio de experto), que puede sesgar los resultados y mermar su reproducibilidad en otros casos de estudio [4]. Este proceso de desagregación se complica más aún si la resolución temporal de las medidas no es suficientemente alta.

Es importante resaltar que en los estudios orientados a clasificar usos finales es habitual trabajar con sensores de alta resolución, que miden los pulsos de consumo por debajo del litro y cada pocos segundos [2]. Este tipo de equipos presentan unas prestaciones muy superiores a los contadores habitualmente disponibles en redes de abastecimiento reales, que miden cada hora o cada pocos minutos en el mejor de los casos. Algunos autores ya han empezado a explorar el efecto de extraer usos finales a partir de series de consumo con una resolución más convencional. Meyer et al [5] definen como “datos rudimentarios” aquellos que se registran en intervalos superiores a 1 minuto o que tienen resoluciones superiores a 0,1 L/pulso. Estas resoluciones, bajas si se comparan con los tradicionales estudios para la caracterización de usos finales, pero altas en comparación con la resolución habitual de las medidas de consumo en redes reales, conllevan retos adicionales derivados de la pérdida de información que se produce al agregar el consumo en intervalos mayores, que enmascaran la contribución relativa de cada uso final [5, 6].

En un esfuerzo por avanzar en esta línea y valorar la capacidad de clasificar usos finales en redes reales, este trabajo pretende clasificar los microconsumos que dan lugar a series agregadas a nivel de vivienda con resolución cada litro y cada minuto. El hecho de trabajar con estas resoluciones, a caballo entre la resolución habitualmente empleada para los estudios de investigación y la propia de sistemas reales, permitirá valorar la viabilidad de extender este tipo de análisis a redes de abastecimiento operacionales.

2. METODOLOGÍA

Este trabajo presenta una metodología que permite identificar y clasificar los microconsumos domésticos de agua potable a partir de señales de consumo por vivienda de resolución por litro y por minuto. Para poder evaluar su eficacia, el método se ha probado en primera instancia con series de consumo simuladas, generadas a partir de un modelo de microconsumos tipo SIMDEUM [7].

2.1 GENERACIÓN DE SERIES DE CONSUMO

Los modelos de microconsumos permiten generar series de consumo a partir de la simulación aleatoria de los usos finales presentes en la vivienda, que son accionados de forma individual por las personas que la habitan. De acuerdo con la formulación presentada en [7], el consumo de agua a nivel de vivienda (en L/s) se puede estimar como:

SMART WATER:

Transición hacia sistemas inteligentes, sostenibles y resilientes

$$Q = \sum_{k=1}^M \sum_{j=1}^N \sum_{i=1}^{F_{jk}} B(I_{ijk}, D_{ijk}, \tau_{ijk}) \left[\frac{L}{S} \right] \quad (1)$$

con

$$B(I_{ijk}, D_{ijk}, \tau_{ijk}) = \begin{cases} I_{ijk} & \text{si } \tau_{ijk} < T < \tau_{ijk} + D_{ijk} \\ 0 & \text{en otro caso} \end{cases} \quad (2)$$

donde k es un contador que se refiere a los M usos finales presentes en la vivienda, j a los N usuarios que viven en ella e i a los tiempos en los que se utiliza cada uso final de acuerdo con su frecuencia F_{jk} . Por su parte, D representa la duración de cada pulso de consumo en segundos, I su intensidad (L/s) y τ_{ijk} al tiempo de inicio de cada consumo. Los parámetros de intensidad, duración y frecuencia (IDF) se generan de forma aleatoria e independiente para cada vivienda de acuerdo con funciones de distribución conocidas. En este trabajo se han utilizado los parámetros presentados en [8] para los siguientes usos finales: grifo de la cocina, grifo del baño, retrete, bañera, ducha, lavavajillas y lavadora. Puesto que se pretende analizar la capacidad de clasificación en series con baja resolución, los caudales obtenidos a partir de las ecuaciones (1)-(2) se han agregado por minuto y redondeado al número entero más bajo para simular consumos por litro y por minuto.

2.2 IDENTIFICACIÓN Y CLASIFICACIÓN DE EVENTOS

El método propuesto para la identificación y clasificación de eventos parte de la serie temporal de consumo en la vivienda, de la que se extrae la serie correspondiente a cada evento para luego transformarla en un dato estructurado con unas características previamente seleccionadas para diferenciar el evento de los demás y así identificar el uso que lo ha generado. A efectos de clasificación, se trabajará con los siguientes usos finales: grifos (sin distinción entre ellos), retretes, bañeras, duchas, lavavajillas y lavadoras. Estos usos se corresponden con tres tipos diferentes de eventos: eventos continuos, que tienen volumen distinto de cero en toda su duración (acotada entre 1 y 3 minutos); eventos continuos independientes, correspondientes a eventos que no tienen valores cero en toda su duración, pero que tienen una duración mínima de 4 minutos; y eventos discontinuos, consistentes en una sucesión de eventos continuos separados por uno o más minutos sin consumo. Estos últimos se contemplan para distinguir las lavadoras y lavavajillas (que suelen concentrar su consumo de agua en momentos o descargas puntuales propias del ciclo de lavado) y los consumos de baja intensidad (que no llegan al litro por minuto y quedan representados en la serie de forma discontinua).

El proceso de identificación y clasificación se compone de cuatro algoritmos o pasos principales:

- Separación y etiquetado de los eventos existentes en la serie temporal. En esta fase se recorren todos los valores de la serie temporal de manera ordenada, teniendo en cuenta el valor y la posición para comprobar si cada volumen es mayor que el valor base (volumen nulo para el evento a aislar). Esta separación permite obtener una lista de eventos que se deben etiquetar. El etiquetado se realiza comparando los valores del evento con el segmento de la serie temporal simulada de cada uno de los posibles usos. Si la serie del evento coincide con algún uso, se etiqueta como tal, y si no, se

SMART WATER:

Transición hacia sistemas inteligentes, sostenibles y resilientes

etiqueta provisionalmente como indefinido. Esta categoría incluirá por defecto todos los eventos en los que se produce el solape de varios usos.

- Análisis interno de eventos. Para reducir el número de eventos indefinidos, se aplica un segundo algoritmo que los analiza internamente. Este proceso consiste en etiquetar minuto a minuto los eventos, agrupando luego las etiquetas adyacentes iguales.
- Agrupación de eventos continuos independientes interrumpidos. Este tercer algoritmo agrupa los eventos continuos independientes interrumpidos de acuerdo con un umbral de proximidad, fijado en este trabajo como 3 minutos (duración máxima de eventos continuos). Si en lugar de trabajar con datos simulados, se trabajase con valores reales, sería necesario plantear un estudio estadístico de la separación entre eventos, ajustar el umbral de proximidad y la forma de agregar los valores de ambos eventos.
- Agrupación de eventos discontinuos. Por último, se agrupan los eventos continuos que forman parte de eventos discontinuos de larga duración, como pueden ser los generados por lavadoras o lavavajillas, durante los cuales ocurren múltiples descargas (eventos continuos). El objetivo de este algoritmo es representar fielmente el evento generado por estos electrodomésticos, en los que la duración es una característica importante para su clasificación.

Los detalles de la metodología propuesta se pueden consultar en [9]. A lo largo de este ciclo de ejecución, se van calculando (y actualizando) las características más definitorias de cada evento. Se utilizan los siguientes atributos: volumen máximo alcanzado durante la duración del evento, volumen mínimo de consumo generado por el evento, minuto de inicio del evento, duración del evento, intensidad media del consumo, número total de minutos en los que no se ha consumido agua, y percentiles 25, 50 y 75 de los volúmenes del evento.

La tabla de eventos resultante constituye la entrada al algoritmo clasificador. La clasificación se realiza mediante la aplicación de un algoritmo de aprendizaje basado en árboles de decisión (CART), que aprovecha la capacidad de explicar y visualizar su comportamiento que ofrece el árbol de decisión resultante para verificar que las características seleccionadas son las más adecuadas.

3. RESULTADOS

Es importante resaltar que el conjunto de datos tiene un reparto desigual entre los distintos usos, ya que la frecuencia de algunos usos (como los grifos o los retretes) es muy superior a la de otros usos (como las bañeras). Esto hace que la eficacia de la clasificación del árbol de decisión se haya evaluado en este trabajo a partir de una métrica de exhaustividad, que se define como la capacidad del modelo para reconocer los eventos de una clase concreta. Esta métrica adquiere un valor 0 si no se clasifica correctamente ningún evento, y se aproxima a 1 a medida que aumenta la tasa de acierto en la clasificación de los eventos.

La Tabla 1 recoge los resultados de exhaustividad para cada tipo de uso final y muestra que, aunque la exhaustividad global alcanza un valor de 0,93, su valor es superior en los eventos continuos asociados a un gran volumen de agua (duchas y llenado de bañeras), más fácilmente distinguibles. La exhaustividad se reduce a 0,76 en los lavavajillas, ya que estos usos se corresponden con eventos discontinuos con una descarga (aproximadamente 3,5 L) similar a la de otros usos, como retretes y grifos.

SMART WATER:

Transición hacia sistemas inteligentes, sostenibles y resilientes

Tabla 1. Exhaustividad de la clasificación por usos finales

Uso final	Exhaustividad
Grifo	0,89
Retrete	0,96
Bañera	1
Ducha	1
Lavavajillas	0,76
Lavadora	1

Estos resultados ponen de manifiesto que, aunque el método es capaz de identificar correctamente los usos de gran volumen, la resolución (por litro y por minuto) limita la capacidad de clasificar usos asociados a volúmenes bajos. También es necesario resaltar que los resultados se han obtenido con datos simulados, y es esperable que la probabilidad de éxito se reduzca cuando se trabaje con datos reales de viviendas de diferente tipo y usuarios con diferentes hábitos.

4. CONCLUSIONES

El trabajo realizado permite probar que las características seleccionadas para definir los eventos son adecuadas y permiten su correcta clasificación a través del árbol de decisión. Estos resultados preliminares en base a series simuladas dejan entrever que este tipo de técnicas de Inteligencia Artificial, en particular las de Aprendizaje Automático o *Machine Learning*, suponen una oportunidad para mejorar el conocimiento sobre el consumo doméstico dentro de cada vivienda, si bien es cierto que todavía es necesario explorar su potencial con datos reales, sometidos a mayor variabilidad.

AGRADECIMIENTOS

Esta publicación es parte del proyecto TED2021-131136B-100 financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR.

REFERENCIAS

- [1] L.A. House-Peters and H. Chang, “Urban water demand modeling: Review of concepts, methods and organizing principles”, *Water Resources Research*, vol. 47, 2011, W05401, <https://doi.org/10.1029/2010WR009624>
- [2] A. Cominola, M., Giuliani, D. Piga, A., Castelletti and A.E. Rizzoli, “Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review”, *Environmental Modelling & Software*, vol. 72, 2015, pp. 198-214, <http://dx.doi.org/10.1016/j.envsoft.2015.07.012>

SMART WATER:

Transición hacia sistemas inteligentes, sostenibles y resilientes

- [3] A. Niknam, H.K., Zare, H., Hosseininasab, A. Mostafaeipour and M. Herrera, “A critical review of short-term water demand forecasting tools – What method should I use?”, *Sustainability*, vol. 14, 2022, 5412, <https://doi.org/10.3390/su14095412>
- [4] F. Mazzoni, S. Alvisi, M. Blokker, S.G. Buchberger, A. Castelletti, A. Cominola, M. Gross, H.E. Jacobs, P. Mayer, D.B. Steffelbauer, R.A. Stewart, A.S. Stillwell, V. Tzatchkov, V. Alcocer Yamanaka and M. Franchini, “Investigating the characteristics of residential end uses of water: A worldwide review”, *Water Research*, vol. 230, 2023, 119500, <https://doi.org/10.1016/j.watres.2022.119500>
- [5] B.E. Meyer, H.E. Jacobs and A. Ilemobade, “Extracting household water use event characteristics from rudimentary data”, *Journal of Water Supply: Research and Technology – AQUA*, vol. 69, no. 4, 2020, pp. 387-397, <https://doi.org/10.2166/aqua.2020.153>
- [6] R. Cardell-Oliver, “Water use signature patterns for analyzing household consumption using medium resolution meter data”, *Water Resources Research*, vol. 49, 2013, pp. 8589-8599, <https://doi.org/10.1002/2013WR014458>
- [7] M. Blokker, J.H.G. Vreeburg and J.C. van Dijk, “Simulating residential water demand with a stochastic end-use model”, *Journal of Water Resources Planning and Management*, vol. 136, no. 1, 2010, pp. 19-26, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000002](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000002)
- [8] M. Blokker, J.H.G. Vreeburg and J.C. van Dijk, “Simulating residential water demand with a stochastic end-use model”, *Journal of Water Resources Planning and Management*, vol. 136, no. 1, 2010, pp. 19-26, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000002](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000002)
- [8] S. Díaz and J. González, “Analytical stochastic microcomponent modeling approach to assess network spatial scale effects in water supply systems”, *Journal of Water Resources Planning and Management*, vol. 146(8), 04020065, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001237](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001237)
- [9] I. García, Sistema de clasificación de microconsumos de agua en viviendas, Trabajo Fin de Grado Universidad de Castilla-La Mancha, Ciudad Real (España), 2023