Watermatex 2015

9th IWA Symposium on Systems Analysis and Integrated Assessment
14-17 June 2015, Gold Coast, Australia

Knowledge discovery from household smart water meters: a customer segmentation case study

Cardell-Oliver, R.*, Wang, J.*
*CRC for Water Sensitive Cities, PO Box 8000, Clayton, Victoria, 3800, Australia and The University of Western Australia, 35 Stirling Highway, Crawley, 6009, Australia

Keywords: smart metering; customer segmentation; data mining

Summary of key findings

Habits are regular, high-magnitude water use activities. Habits account for a high proportion of household demand for water and they are easily recognisable by humans. We propose an automatic algorithm for discovering habits in hourly smart water meter time series. Customers can be segmented based on the characteristics of their habits. Using two real-world case studies, we demonstrate that customer segmentation based on habits yields valuable insights for water providers and their customers.

Background and relevance

Drinking water is a valuable resource. Decreasing stream flows, rising populations, and increasing costs of operations and infrastructure, mean that water efficiency and demand management are key challenges for the industry to achieve its triple bottom line of economic, environmental and societal goals (NWC 2014, Atkinson and Medbury 2013).

Customer segmentation aims to identify customers with similar behaviours. Specific groups of customers can be targeted for business use cases in efficiency, planning, or marketing. In the electricity industry, customer segmentation is widely used to design load-shifting strategies (Albert and Rajagopal 2013, Kwac et al 2014). In the water industry, segmentation can be used to identify customers with inefficient behaviours such as leaks in apartment buildings (Kermany et al 2013). Another water application is load shifting of discretionary demand away from the peak hour of day. Knowledge discovery from smart water meter data can be used for customer segmentation based on evidence of customers' actual water consumption.

Previous studies have used hourly demand thresholds to identify leaks (Britton et al 2008), probable outdoor activity (Cole et al 2013) and other water use signatures (Cardell-Oliver 2013) in hourly smart meter data. Recently, algorithms have been proposed for discovering regular high magnitude (RHM) behaviours that identify customers' recurrent habits or routines (Cardell-Oliver 2014, Wang et al 2015). Habits and routines are known to be important in the psychology of water use behaviour (Russell and Fielding 2010). RHM behaviours (i.e. habits or routines) account for a high proportion of all demand and they are easily recognisable for humans.

We introduce novel customer segmentation methods based on recurrent, high-magnitude (RHM) behaviours. To the best of our knowledge this is the first study to propose customer segmentation methods based on behaviours discovered from household smart meter data.

Results

Our methods for customer segmentation are based on two complementary types of RHM behaviours: habits and routines. Habits are clusters of smart meter observations that have similar volume, occur at the same hour of day, and have a persistent, regular recurrence pattern (eg. every day or every Mon, Wed and Fri for at least 4 weeks) (Cardell-Oliver 2014). Habits capture human water-use behaviours such as automatic watering systems programmed to run at a fixed hour on certain days. Routines are short, similar sequences of consecutive hours of demand that occur frequently (Wang et al 2015). The consumption in each routine sequence must include at least one high magnitude hour. But routines need not necessarily occur at the same time of day or on a regular pattern of days. Routines capture human water-use behaviours such as hand watering a garden or weekly washing a car that recur frequently but are time-flexible.

Two customer segmentation methods are proposed: temporal and schedule segmentation. Temporal segmentation partitions users according to the timing of their RHM behaviours. Water demand for each hour of the day is partitioned into RHM and non-RHM use, and the RHM use is further partitioned into the highest consuming 5% (say) of customers, and the rest. From this, we can identify RHM contribution to peak hour of day, as well as the peak RHM hour of day, which may be different. Opportunities for demand shifting can also be identified if the peak hour and its RHM contribution are substantially higher than other demand hours.

Schedule segmentation partitions users according to the frequency and intensity of their RHM behaviours. The data used for this segmentation is all habits identified for the population. Each habit occurs at a particular time of day on a specific pattern of days. For example, FOR mt=98 ACTIVITY apx. 1511 L/h OCCURS Mo, We, Fr, Su- day AT 6:00 FROM 16 Mar'14 to 6 Apr'14. The intensity of a habit refers to its hourly rate, in this example it is apx. 1511 L/h. The frequency of a habit refers to the number of times it occurs per week, in the example 4 times a week. The significance of a habit is the aggregate of all water used during the life of the habit. The example has a significance of 19.9~kL (13 x 1511 L). For schedule segmentation, each habit is mapped to one frequency-intensity category.

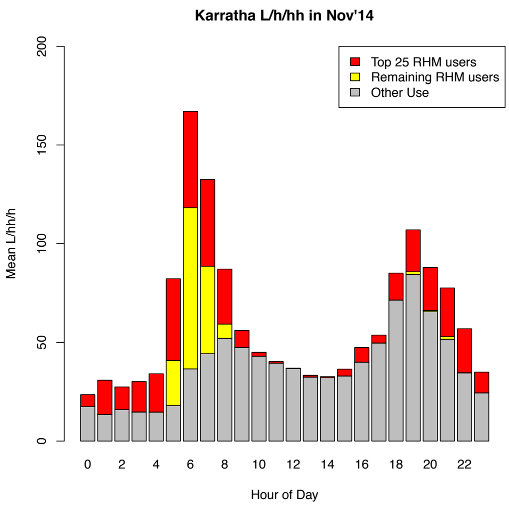

Figure 1.1 Segmented hourly demand for Karratha in November (summer)

The proposed temporal and schedule segmentation methods are evaluated using smart water meter data for 1000 customers over one year from two contrasting populations. Both populations are in climate zones that are challenging for water conservation: Karratha in the coastal North West of Western Australia, and Kalgoorlie in the inland Eastern goldfields of Western Australia. Using 7 business use cases, we show that customer segmentation on RHM behaviours can identify significant savings. The considered cases are:

Q1. When are the peak demand hours of the day, and which RHM users contribute most to this?
Q2. When are the peak RHM hours, and which RHM users contribute most to this?
Q3. When are the outlier peak hours with high RHM activity, and which users contribute most to this?
Q4. Which users have the most frequent and most intense RHM behaviours?
Q5. Which users have high intensity watering schedules of higher frequency than the water restriction rosters?
Q6. Which users and which frequency-intensity classes of RHM behaviour account for the highest proportion of consumption?
Q7. Are the users with high RHM behaviours the same as those with high overall use?

For the case study samples, time-flexible RHM routines accounted for 23% of consumption for Kalgoorlie in ABCB climate zone[1] 4, but time-fixed habits (e.g. from automatic watering systems) for only 13%. Karratha, in climate zone 1, had 45% of demand as time-flexible routines and 34% as time-fixed habits. RHM behaviours were identified for 63% to 85% of households.

Discussion

Both habits and routines have distinctive but population-specific time-of-day distributions. Temporal segmentation can be used to identify customers in order to reduce demand at relevant times of the day or alternatively for shifting peak demand. Figure 1 shows the segmentation of hourly demand in Karratha for November (i.e. summer) segmented into consumption of the top 5% of RHM use (red), remaining RHM use (yellow) and non-RHM use (grey). Peak demand at 6am is dominated by RHM behaviours. The red customer segment at 6am (top 5% of users) could be targeted to reduce demand at 6am or alternatively for shifting peak demand.

Schedule segmentation was used to identify different styles of water use categorised by the frequency and intensity of the behaviour. Each population had different distributions of these behaviours. The most significant schedule for Karratha, by total consumption percentage, was every-day at 500-100 L/h. This schedule is not consistent with local watering restrictions of alternate days. For Kalgoorlie, the most significant schedule was
twice-a-week patterns at above 1000 L/h. This schedule is consistent with local watering restrictions, but the high intensity of these events may indicate inefficient watering.

In future work we plan to investigate new segmentation methods that incorporate contextual information such as locality, lot size and house ownership. We are also investigating knowledge discovery techniques for detecting anomalies in smart meter data and for quantifying leak anomalies in water pipeline distribution systems.

References

A. Albert and R. Rajagopal. Smart meter driven segmentation: What your consumption says about you. *IEEE Transactions on Power Systems*, 28(4):4019–4030, Nov 2013.

S. Atkinson and K. Medbury. Driving water efficiency improvements in large customers: linking customer engagement and high resolution water use data. In *AWA Water Efficiency Conference*, Brighton Beach, Sydney. Australian Water Association, March 2013.

T.C. Britton, G. Cole, R.A. Stewart, and D. Wiskar. Remote diagnosis of leakage in residential households. *Water: Journal of the Australian Water Association*, 35(6):56–60, September 2008.

R. Cardell-Oliver. Water use signature patterns for analyzing household consumption using medium resolution meter data. *Water Resources Research*, pages 1–11, 2013.

R. Cardell-Oliver. A habit discovery algorithm for mining temporal recurrence patterns in metered consumption data. In *1st International Workshop on Machine Learning for Urban Sensor Data (SenseML)* 15 Sep 2014. www.tk.informatik.tu-darmstadt.de/de/senseml-2014/

G. Cole and R. A. Stewart. Smart meter enabled disaggregation of urban peak water demand: precursor to effective urban water planning. *Urban Water Journal*, 10(3):174–194, 2013.

E. Kermany, H. Mazzawi, D. Baras, Y. Naveh, and H. Michaelis. Analysis of advanced meter infrastructure data of water consumption in apartment buildings. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '13*, pages 1159–1167, New York, NY, USA, 2013. ACM.

J. Kwac, J. Flora, and R. Rajagopal. Household energy consumption segmentation using hourly data. *IEEE Transactions on Smart Grid,* 5(1):420–430, Jan 2014.

[1] http://www.abcb.gov.au/major-initiatives/energy-efficiency/climate-zone-maps

NWC. National performance report 2012-13: urban water utilities. *National Water Commission,* 2014. http://www.nwc.gov.au, accessed [7 January 2015].

S. Russell and K. Fielding. Water demand management research: A psychological perspective. *Water Resources Research*, 46(5), 2010.

J. Wang, R. Cardell-Oliver, and W. Liu. Efficient discovery of recurrent routine behaviours in smart meter time series by growing subsequences. In *PacificAsia Conference on Knowlwedge Discovery in Databases*, Ho Chi Minh City, Vietnam, May 2015.