Leakage Detection With Modified Decomposition Models

Sverrir Heiðar Davíðsson, Sigríður Sigurðardóttir, Sverrir Guðmundsson
Veitur Utilities PLC

INTRODUCTION

Non-revenue water (NRW) is difference between water put into water distribution systems (WDS) and the water billed to consumers. About 80% of NRW is estimated to be caused by leakages from pipes. The total cost of NRW worldwide is estimated to be more than $14 billion per year. [3]

There are many challenges in reducing leakages, but three core challenges are: detection, size estimation and localization. Detection and localization are strictly necessary in order to repair leakages. Size estimation opens the possibility for prioritization by size of leakage and cost of repairs which can lead to a more efficient and effective reduction of leakages.

The method described here is mainly one of size estimation, but detection can be achieved by monitoring for changes in total leakage amount estimated in the system. The method can only tell that a leakage is present downstream of the meter but cannot estimate its location any more precisely. This method is applied to time series data from a flow meter and can estimate how much of the total demand downstream from the meter stems from leakages. The method is a type of semi-blind source separation using a decomposition model known as NMF and has been modified specifically for leakage estimation for water utilities.

CASE STUDY : SKERJAFJÖRÐUR

The method is applied here to 4 years of data (See Figure 2 below) from a flow sensor that measures water input into a small residential neighbourhood (See Figure 3 below) in Reykjavík, Iceland.

- The modified NMF model is trained on this data to learn the patterns
- A pattern length of 1 week is used
- The abundance of each pattern for each week is estimated to minimize reconstruction error
- Known leakage repairs are used as proxies for leakages when evaluating the results.

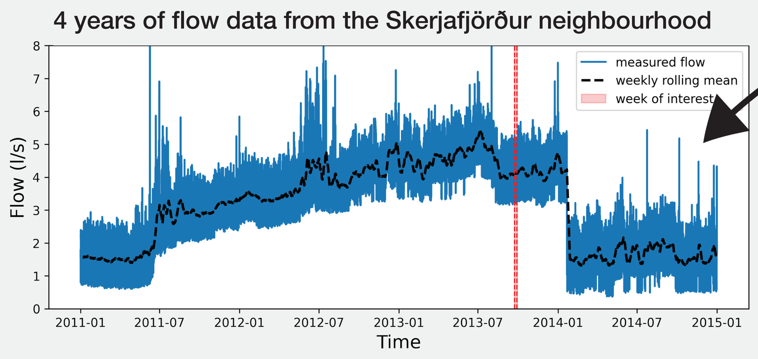

Figure 2 – Four years of flow data from the Skerjafjörður neighbourhood. Highlighted between red lines is a single week for further examination in Figure

Figure 3 – Location of the flow sensor for the Skerjafjörður neighbourhood from which the data on Figure 2 comes.

DISCUSSION

Topics for further study:
- Performance in non-residential areas
- Performance on higher or lower-resolution time series data
- Analysis of non-leakage sources of flat usages
- Deciding optimal number of patterns to search for in the data

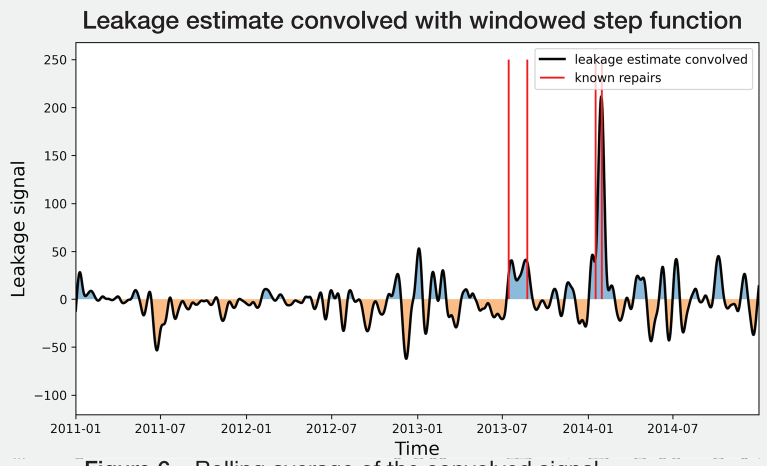

Figure 6 – Rolling average of the convolved signal.
Positive values indicate decrease in leakage and vice versa

Leakage "detection"
In order to automate the detection of leakages with this method there needs to be something to trigger a detected leakage, such as a sudden increase in the leakage estimate.

One way to detect increases in a signal is by convolving the signal with a windowed step function. The result of such an operation is shown in figure 6.

[1] de Almeida, André & Favier, Gerard & Carvalho Lustosa da Costa, Joao Paulo & Mota, joão. 2016. Overview of Tensor Decompositions with Applications to Communications. SIAM Review, (Aug. 2009), 455–500. DOI:https://doi.org/10.1137/07070111X

[2] Cichocki, Andrzej and Phan, Anh-Huy. 2009. Fast Local Algorithms for Large Scale Nonnegative Matrix and Tensor Factorizations. IEICE Transactions. 92-A. 708-721. DOI:http://doi.org/10.1587/transfun.E92.A.708

[3] Kingdom, Bill; Liemberger, Roland; Marin, Philippe.2006. The challenge of reducing non-revenue water (NRW) in developing countries - how the private sector can help : a look at performance-based service contracting (English). Water Supply and Sanitation Sector Board discussion paper series no. 8; Retrieved November 30, 2020 from http://documents.worldbank.org/curated/en/385761468330326484/The-challenge-of-reducing-non-revenue-water-NRW-in-developing-countries-how-the-private-sector-can-help-a-look-at-performance-based-service-contracting

[4] Scikit-learn, 2020. sklearn.decomposition.NMF — scikit-learn 0.23.2 documentation. Retrieved November 30, 2020 from https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.NMF.html

PROBLEM AND METHOD

Description of the problem (See figure 1)
- Assume 3 types of users in the system
- Each user type has a water usage pattern
- The patterns are:
  - Non-negative (i.e. ≥0)
  - Unknown to us
  - Fixed over time
- The number of each user type:
  - Is non-negative (i.e. ≥0)
  - Is unknown to us
  - Is not fixed over time
- We only observe the combined flow
- The objective is to find the patterns and the number of each type of user for each day

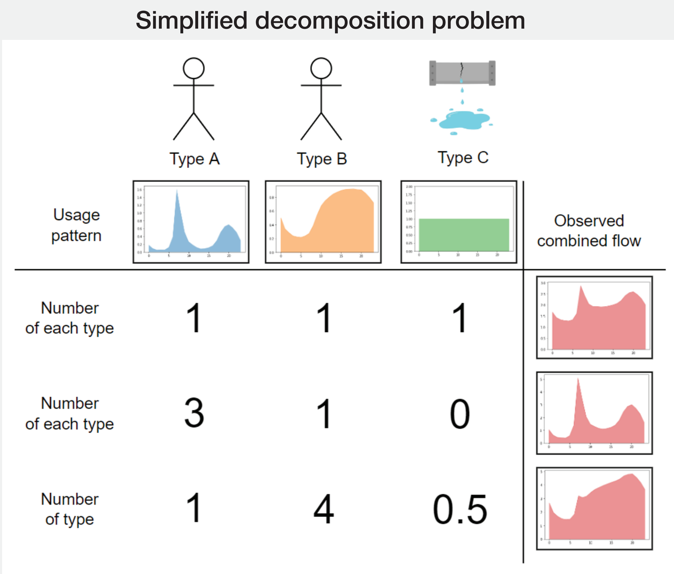

Figure 1 – A simplified decomposition problem

Description of the method
This problem can be solved with Non-negative Matrix Factorization (NMF) [4] if we treat each day as a 24-dimensional vector with the values in the vector corresponding to flow observations at a given time of day.

The approach described here uses a modified NMF model to separate different patterns in the system. The modification is the additional constraint that one of the patterns must be flat. This constraint is imposed during the optimization of the model by changing a step in the ALS algorithm [1,2] such that the flat pattern is not updated with the other patterns.

RESULTS

The observed flow for a given week can be represented as the sum of three patterns (See figure 4):
- **Pattern A** (Green)
- **Pattern B** (Orange)
- **Pattern C – leakage pattern** (Blue)

Patterns A and B are learned from the original time series and are considered the best patterns to approximate each week in the original time series.

Note: One possible interpretation of Pattern A is that it represents water usage of people who work 8-16 jobs:
- Peak in mornings on workdays
- Even usage during weekends

Pattern B may be interpreted as usage of people who stay at home.

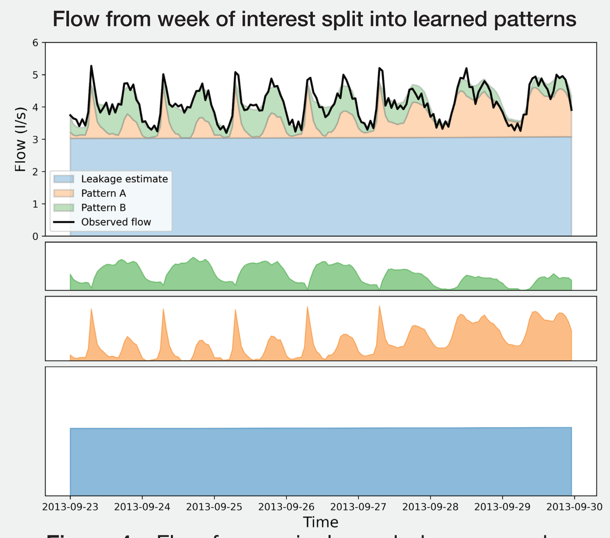

Figure 4 – Flow from a single week decomposed into its constituent patterns

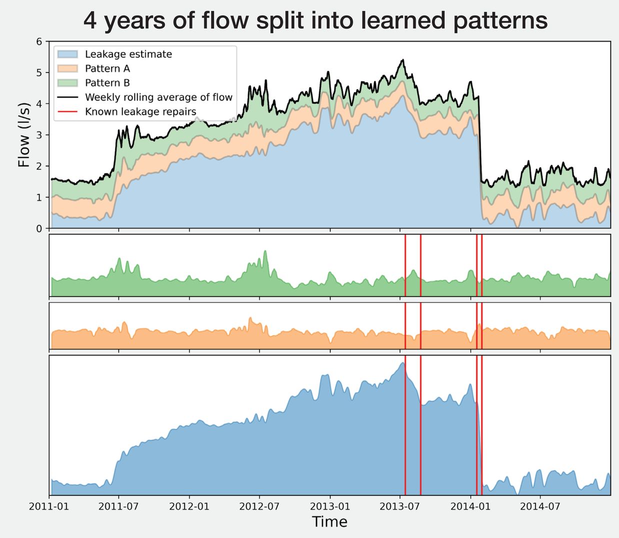

Figure 5 – Rolling average of flow over the dataset and the abundance of each constituent pattern across the period

The graph shows the estimated abundance of each pattern across the four-year period.

In the three lower plots we see red lines that indicate times where four known leakage repairs took place.

The effect of these repairs is only obvious in the leakage estimate. The others appear unaltered.

CONCLUSIONS

We proposed the use of a modified NMF for leakage estimation and evaluated the quality of the method by using leakage repairs as proxies for leakages.

The results showed a great decline in estimated leakages following each of the known repairs while other estimated usage remained stable following these repairs.

The results indicate that the method can successfully isolate the leakages from other usage when applied to data of this nature.

VEITUR

the international water association

Inspiring change

www.iwahq.org