Exploring unusual sensor behaviour in buildings using BMS data and unsupervised learning techniques

Cameron Roach (PhD candidate)

31 July 2018

Overview

1. Motivation
2. Feature engineering
3. Unsupervised learning techniques
4. Examples
5. Thoughts and reflections

Motivation

- The bigger the BMS the harder it is to find what matters. Locating problems is difficult and time-consuming.
- Help engineers explore unusual BMS behaviour.
- Compare and learn from multiple buildings.

Motivation

Need an easy way to visualise and explore data.

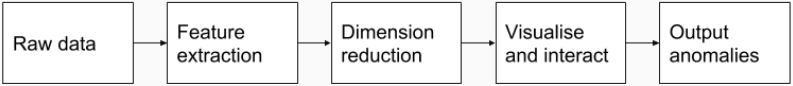
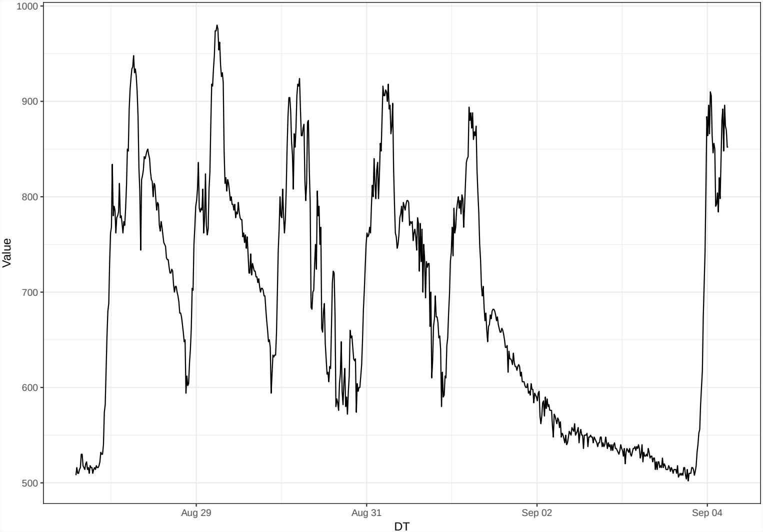
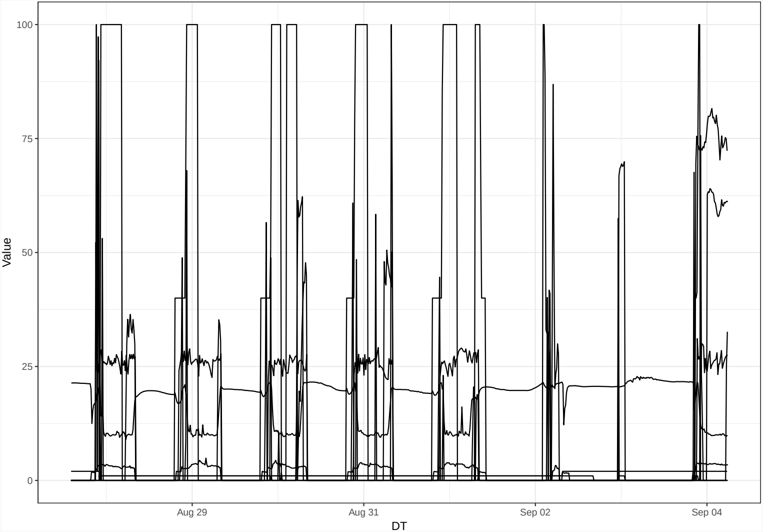


Figure 1: Intended workflow

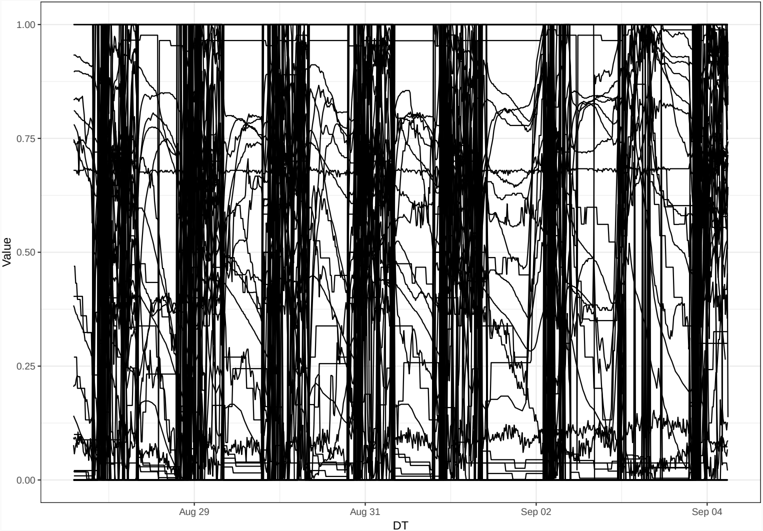
Motivation



Motivation



Motivation



Feature engineering

Data

Focus on 597 AHU points from three separate buildings.

Table 1: Number of points for each measurement type.

Measure	BID0025	BID0126	BID1701
Cooling control valve (CCV)	20	33	6
Economy cycle dampers (ECD)	18	33	6
Enabled (ENB)	22	34	8
Return air temperature (RAT)	1	32	2
Supply air pressure (SAPR)	16	30	8
Supply air pressure setpoint (SAPRSP)	16	29	8
Supply air temperature (SAT)	20	32	8
Supply air temperature setpoint (SATSP)	16	32	6
Speed (SPD)	16	30	8
Status (STS)	20	33	8
VAV damper position max (VAVDM)	16	30	0

Feature engineering

Metadata features

- Buildings have inconsistent point descriptions. However there are often some useful acronyms hidden within the names.
- Character-level bigrams and trigrams are created for each BMS point's name. Whitespace and numeric values are omitted.
- For example, the first four bigrams of "NAE-08/FC-1.FD-88 AHU-14-1.AHU-14-1 VAV DMPR-POS" will be na, ae, ef and fc. The first four trigrams are nae, aef, efc and fcf.

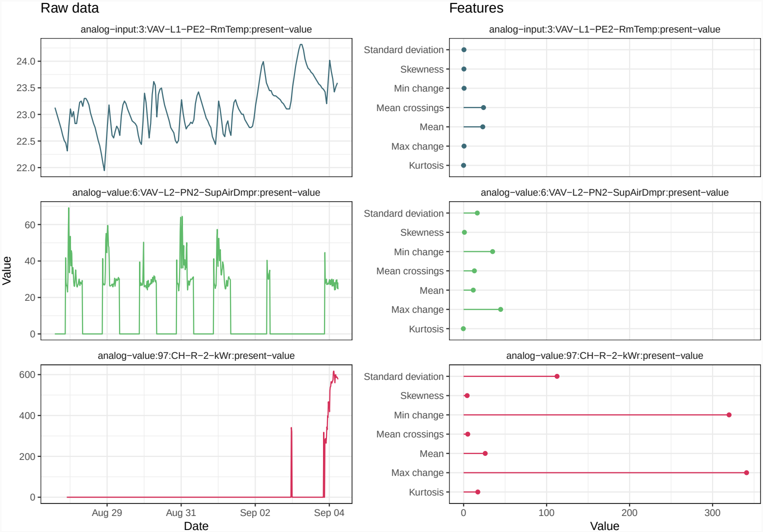
Feature engineering

Time series features

Time series sampled at irregular intervals. Instead of interpolating raw time series (which can corrupt signal) we engineer global features that describe entire time series. These time series features include:

- Number of unique values
- Mean
- Maximum value
- Minimum value
- Standard deviation
- Skew
- Kurtosis
- Maximum change
- Minimum change
- Number of mean crossings.

Feature engineering



Unsupervised learning techniques

Approach the problem using dimensionality reduction and clustering.

Linear projections:

- principal component analysis
- sparse principal component analysis

Non-linear projections (manifold learning):

- isometric mapping
- t-distributed stochastic neighbour embedding
- spectral embedding (nearest neighbours affinity matrix)
- spectral embedding (radial basis function affinity matrix).

Principal component analysis

- PCA is an unsupervised learning technique that has been used in various fault detection approaches.
- Despite its popularity it does have some drawbacks that need to be considered.
- PCA focuses on producing orthogonal components that capture as much variation in the data as possible.

Drawbacks:

- Does not aim to preserve proximity relationships between points and neighbourhoods.

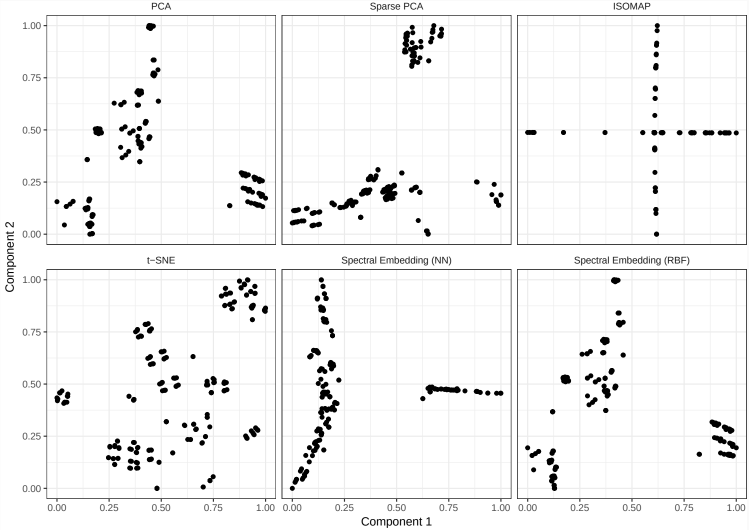
t-distributed stochastic neighbour embedding

- t-SNE attempts to preserve nearest neighbours.
- Well suited to visualising high dimensional spaces in two or three dimensions as it plots similar objects nearby and dissimilar objects far away with high probability.

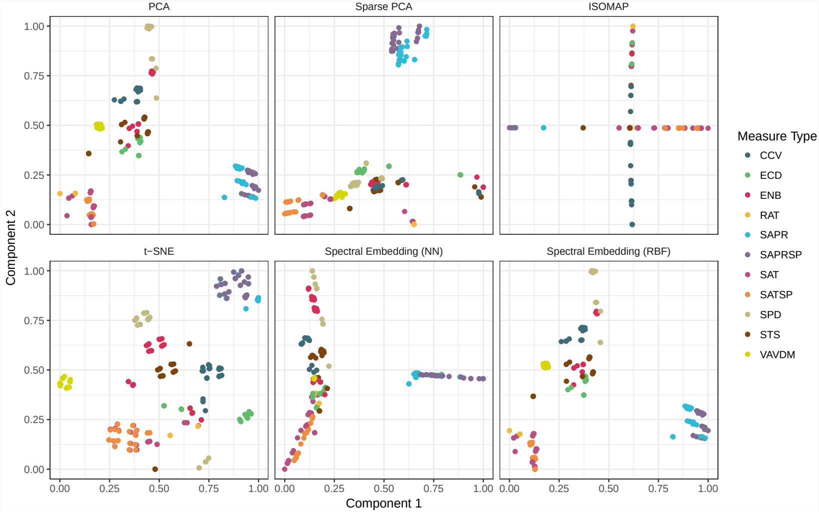
Drawbacks:

- $T_{t-SNE}(N) = \mathcal{O}(N \log(N))$ (Barnes-Hut) whereas $T_{PCA}(N) = \mathcal{O}(N)$.
- Doesn't preserve distance or density. Don't use distance or density based clustering algorithms after t-SNE!

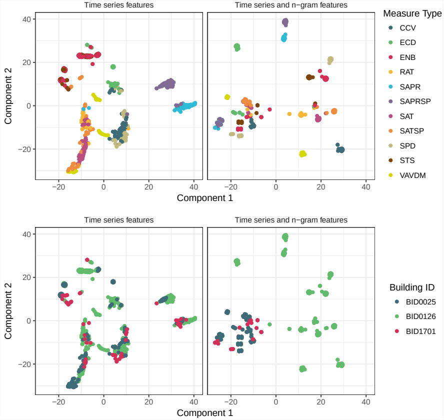
Dimensionality reduction



Dimensionality reduction



Dimensionality reduction



- Different naming conventions between buildings can cause issues.
- Clusters may represent buildings rather than sensor types.
- Possible that only using time series features may be better.

Examples

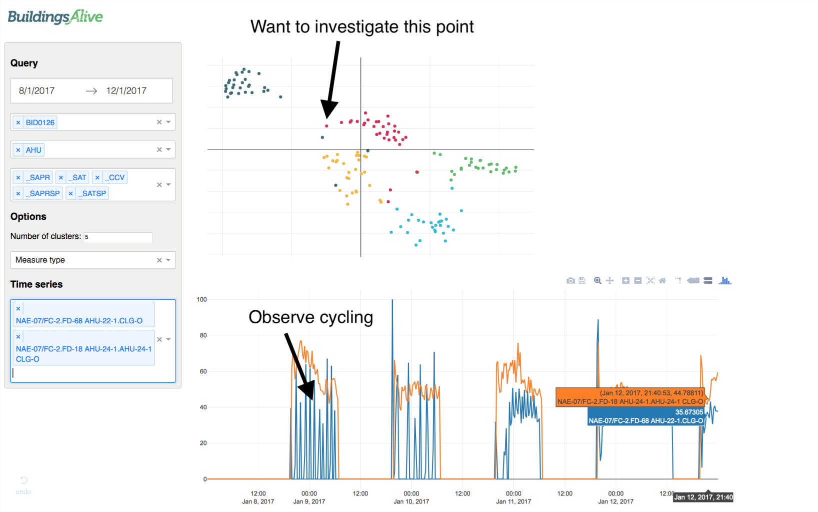


Figure 2: Control valve short cycling.

Examples

Figure 3: Cooling control valve only open for two hours.

Thoughts and reflections

- Futher work needs to be done selecting appropriate metadata and time series features.
- Introduce calculated time series, e.g., SAT - SATSP. View relationships between points.
- Instead of comparing points, compare point histories day by day.
- Need to pick the right tool for the job when visualising or clustering data.
- Speeds up the process of exploring BMS data sets.

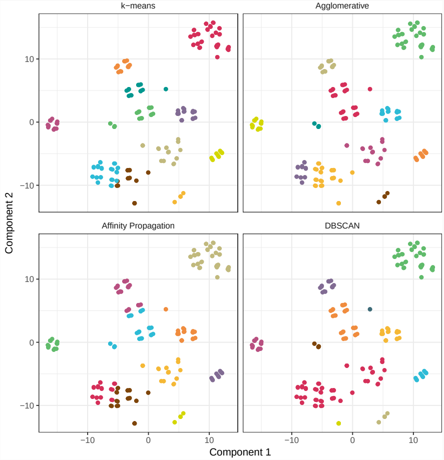
Questions?

Clustering

Clustering was performed on lower dimensional space produced by t-SNE. Again, several methods were tested:

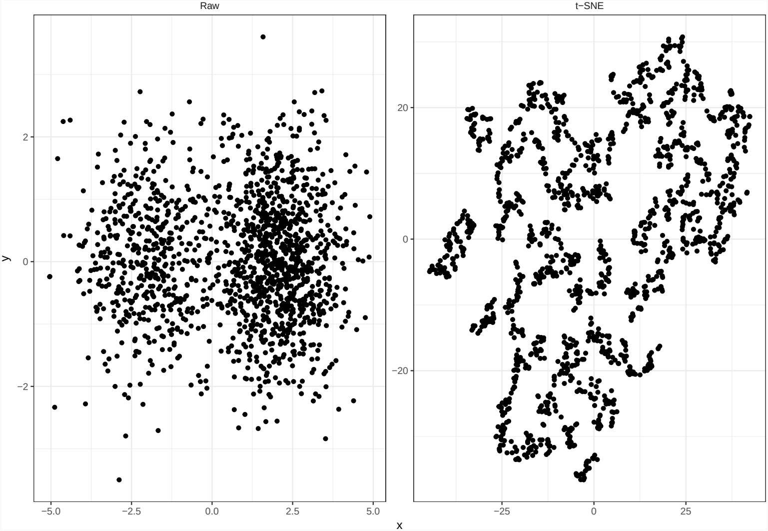
- k-means
- agglomerative clustering
- affinity propagation
- DBSCAN.

Clustering



- Each clustering algorithm appears to offer similar performance.
- Possibly because t-SNE produces clear separation between clusters.
- Clustering on t-SNE is risky and so visualisation and exploration is preferred.

Clustering after t-SNE



Clustering after t-SNE