university of groningen

faculty of mathematics and natural sciences

Mining household appliances patterns by monitoring electric plug loads

Master Thesis Computing Science

November 22, 2016

Student: ing. Mathieu Kalksma

Primary supervisor: Prof. dr. ir. Marco Aiello

Secondary supervisor: Prof. dr. Alexander Lazovik

ABSTRACT

About 20% of the total energy consumption in the Netherlands is consumed by households. For this thesis an attempt is made to mine patterns for household appliances by monitoring electric plug loads which later can be used to reduce the energy consumed. Three experiments are conducted on data sets from three buildings: two households and one office building. For the experiments Markov models are created which characterise the usage of household appliances over time. Experiments are conducted to predict the usage of household appliances in the near future, an experiment to recover erroneous sensor readings and finally an experiment to forecast the upcoming energy consumption in real-time. Usage prediction is predicted correctly for over 87% for all the data sets while the recovery of erroneous sensor results in the same prediction for at least 80% of the time. The accuracy for forecasting the energy consumption of devices depends highly on the device with the Mean Absolute Percentage Error (MAPE) varying from 6% to 70%.

ACKNOWLEDGEMENTS

I would like to express my gratitude to everyone who have made it possible for me to complete this thesis. First of all I would like to thank my supervisors *Marco Aiello* and *Alexander Lazovik* for their professional and accurate support they have given me during the research.

Further I want to express my gratitude especially to PhD student *Brian Setz* for providing me with data from experiments conducted at the University of Groningen and his help with the data.

As last I want to thank my fellow students who were always up for a discussion about the meaning of data or best approaches to solve an issue.

CONTENTS

1 INTRODUCTION 7
1.1 Research question 8
1.2 Document structure 8

2 MODEL DESIGN AND ALGORITHMS 9
2.1 Model 9
2.1.1 Preliminaries 10
2.1.2 Training algorithm 11
2.2 Device state prediction 13
2.3 Energy consumption prediction 14
2.4 Fault tolerance prediction 15

3 DATA SETS 17
3.1 RUG data set 17
3.2 Electronic Consumption Occupancy data set 17
3.2.1 Household 1 18
3.2.2 Household 2 18

4 EXPERIMENT ONE: STATE PREDICTION 19
4.1 Parameter tweaking 19
4.1.1 Minimal support threshold r 19
4.1.2 Sequence length k 20
4.1.3 Time-unit lenth m 21
4.1.4 Predictor method 22
4.1.5 Optimal configuration 23
4.2 Results 23
4.2.1 RUG data set 23
4.2.2 ECO data set household 1 24
4.2.3 ECO data set household 2 25
4.3 Conclusion 26

5 EXPERIMENT TWO: ENERGY CONSUMPTION PREDICTION 27
5.1 RUG data set 27
5.2 ECO data set household 1 28
5.3 ECO data set household 2 29
5.4 Conclusion 30

6 EXPERIMENT THREE: FAULT TOLERANCE 31
6.1 RUG data set 31
6.2 Household 1 32
6.3 Household 2 33
6.4 Conclusion 34

7 RELATED WORK 37
7.1 State prediction 37
7.2 Energy consumption 39
7.3 Fault tolerance 40
7.4 Other usages of Markov models 41

8 CONCLUSION AND FUTURE WORK 43
8.1 Future work 44

A DETAILED RESULTS FOR STATE PREDICTION 45
A.1 RUG data set 45
A.1.1 Supported sequences per device 45
A.1.2 Device specific validations 46
A.2 Detailed results for household 1 52
A.2.1 Supported sequences per device 52
A.2.2 Device specific validations 53
A.3 Detailed results for household 2 56
A.3.1 Supported sequences per device 56
A.3.2 Device specific validations 58

B DEVICE SPECIFIC ENERGY CONSUMPTION PATTERNS 65
B.1 RUG data set 65
B.2 Household 1 67
B.3 Household 2 68

C DETAILED FAULT TOLERANCE RESULTS 71
C.1 Household 1 71
C.2 Household 2 72

Figure 1 Energy consumption distribution in households in the U.S in 1993 and 2009 [4] 7
Figure 2 Example of a Markov model 8
Figure 3 Example of a model with two devices 10
Figure 4 Illustration of a transactions time line 13
Figure 5 Floor plan of the coffee corner in the Bernoulliborg building 17
Figure 6 Effect of the minimal support threshold on # supported sequences. Mined with $|\mathcal{T}| = 131386$, $|V| = 5$ on the RUG data set 19
Figure 7 Effect of support sequence on accuracy. ($r = 0.3$, $m = 10$) 20
Figure 8 Effect of maximum sequence length ($m = 10$, $r = 0.3$, average) 21
Figure 9 Effect of time length ($k = 5$, $r = 0.2$, average) 22
Figure 10 Effect of predictor method ($r = 0.2, k = 5$) 22
Figure 11 Certainty of state predictions on the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 24
Figure 12 Certainty of state predictions on household 1 ($r = 0.2, k = 5, m = 5, wavg$) 25
Figure 13 Certainty of state predictions on household 2 ($r = 0.4, k = 5, m = 5, wavg$) 26
Figure 14 Energy consumption prediction for RUG data set 28
Figure 15 Energy consumption prediction for household 1 29
Figure 16 Energy consumption prediction for household 2 30
Figure 17 Fault tolerance certainty of predictions on the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 32
Figure 18 Fault tolerance certainty of predictions on household 1 ($r = 0.2, k = 5, m = 5, wavg$) 33
Figure 19 Fault tolerance certainty of predictions on household 2 ($r = 0.2, k = 5, m = 5, wavg$) 34
Figure 20 Histogram of certainty of predictions for the microwave in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 48
Figure 21 Histogram of certainty of predictions for the printer in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 49
Figure 22 Histogram of certainty of predictions for the boiler in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 50
Figure 23 Histogram of certainty of predictions for the screen in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 51
Figure 24 Histogram of certainty of predictions for the coffee machine in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$) 52
Figure 25 Certainty of predictions for the fridge in ECO household 1 ($r = 0.2, k = 5, m = 5, wavg$) 54
Figure 26 Histogram of certainty of predictions for the freezer ($r = 0.2, k = 5, m = 5, wavg$) 55
Figure 27 Histogram of certainty of predictions for the washing machine ($r = 0.2, k = 5, m = 5, wavg$) 55
Figure 28 Histogram of certainty of predictions for the dryer ($r = 0.2, k = 5, m = 5, wavg$) 56
Figure 29 Histogram of certainty of predictions for the dishwasher in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 59
Figure 30 Histogram of certainty of predictions for the air exhaust in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 60

Figure 31 Histogram of certainty of predictions for the fridge in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 60
Figure 32 Histogram of certainty of predictions for the freezer in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 61
Figure 33 Histogram of certainty of predictions for the lamp in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 62
Figure 34 Histogram of certainty of predictions for the TV in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 62
Figure 35 Histogram of certainty of predictions for the stereo in household 2 ($r = 0.4, k = 5, m = 5, wavg$) 63
Figure 36 Energy consumption prediction for the screen in RUG data set 65
Figure 37 Energy consumption prediction for the microwave in RUG data set 65
Figure 38 Energy consumption prediction for the boiler in RUG data set 66
Figure 39 Energy consumption prediction for the coffee maker in RUG data set 66
Figure 40 Energy consumption prediction for the printer in RUG data set 66
Figure 41 Energy consumption prediction for the fridge in household 1 67
Figure 42 Energy consumption prediction for the dryer in household 1 67
Figure 43 Energy consumption prediction for the freezer in household 1 67
Figure 44 Energy consumption prediction for the washing machine in household 1 68
Figure 45 Energy consumption prediction for the dishwasher in household 2 68
Figure 46 Energy consumption prediction for the air exhaust in household 2 68
Figure 47 Energy consumption prediction for the fridge in household 2 69
Figure 48 Energy consumption prediction for the freezer in household 2 69
Figure 49 Energy consumption prediction for the lamp in household 2 69
Figure 50 Energy consumption prediction for the TV in household 2 70
Figure 51 Energy consumption prediction for the stereo in household 2 70

LIST OF TABLES

Table 1 Comparison of the different data sets 23
Table 2 Validation of state predictions on the RUG data set 23
Table 3 Validation of state prediction on household 1 of the ECO data set 24
Table 4 Validation of state prediction on household 2 of the ECO data set 25
Table 5 Results of energy prediction on the RUG data set 28
Table 6 Results of energy consumption prediction on household 1 28
Table 7 Results of energy prediction on household 2 29
Table 8 Fault tolerance comparison of the different data sets 31
Table 9 Fault tolerance results for the rug data set 32
Table 10 Fault tolerance results for household 1 33
Table 11 Fault tolerance results for household 2 34
Table 12 Accuracy for Barbato et al on household 2 38
Table 13 Comparison of different energy prediction methods from the literature 40
Table 14 Top 5 supported sequences for the TV-screen in the RUG data set 45
Table 15 Top 5 supported sequences for the microwave in the RUG data set 46
Table 16 Top 5 supported sequences for the boiler in the RUG data set 46
Table 17 Top 5 supported sequences for the printer in the RUG data set 46
Table 18 Top 5 supported sequences for the coffee maker in the RUG data set 47
Table 19 State prediction results for the RUG data set 47
Table 20 Top 5 supported sequences for the fridge in household 1 52
Table 21 Top 5 supported sequences for the freezer in household 1 53
Table 22 Top 5 supported sequences for the washing machine in household 1 53
Table 23 Top 5 supported sequences for the dryer in household 1 53
Table 24 State prediction results for household 1 54
Table 25 Top 5 supported sequences for the dishwasher in household 2 56
Table 26 Top 5 supported sequences for the air exhaust in household 2 57
Table 27 Top 5 supported sequences for the fridge in household 2 57
Table 28 Top 5 supported sequences for the freezer in household 2 57
Table 29 Top 5 supported sequences for the lamp in household 2 58
Table 30 Top 5 supported sequences for the TV in household 2 58
Table 31 Top 5 supported sequences for the stereo in household 2 58
Table 32 State prediction results for household 2 59
Table 33 Top 5 supported fault tolerance sequences for the fridge in household 1 71

Table 34 Fault tolerance results for the fridge in household 1 71
Table 35 Top 5 supported fault tolerance sequences for the freezer in household 1 72
Table 36 Fault tolerance results for the freezer in household 1 72
Table 37 Top 5 supported fault tolerance sequences for the fridge in household 2 72
Table 38 Fault tolerance results for the fridge in household 2 73
Table 39 Top 5 supported fault tolerance sequences for the freezer in household 2 73
Table 40 Fault tolerance results for the freezer in household 2 73
Table 41 Top 5 supported fault tolerance sequences for the TV in household 2 74
Table 42 Fault tolerance results for the TV in household 2 74
Table 43 Top 5 supported fault tolerance sequences for the stereo in household 2 75
Table 44 Fault tolerance results for the stereo in household 2 75

1

INTRODUCTION

In 2015 a Dutch court has ruled that the Dutch government has to put more effort to reduce domestic greenhouse-gas emissions [5]. One way to reduce the emission of greenhouse-gas is to reduce the consumption of energy. There are a lot of approaches to reduce the energy consumption, from large scale as closing power plants [2] to small scale by reducing the energy consumption in a single household. According to the Dutch CBS [1] about 20% of all the energy consumption in the Netherlands is consumed by households [15].

According to the U.S Energy Information Administration [2] the energy consumption for household appliances, electronics and lighting uses about 35% of the total energy consumed in a household. This is an increase of 10% compared to 1993 as shown in Figure 1.

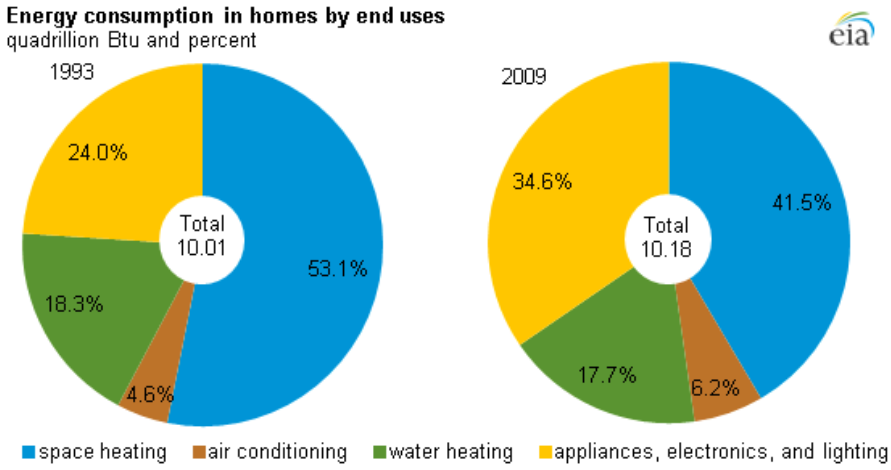

Figure 1: Energy consumption distribution in households in the U.S in 1993 and 2009 [4]

As buildings more and more are being equipped with sensors it is possible to monitor the state of the building. Sensors that monitor the energy consumption on a single socket allow us to monitor an appliance his power consumption in (near) real-time. Meanwhile occupancy sensors give us insight in the number of people in a room and light sensors tell us if a room is illuminated or not.

In this research we try to mine patterns for household appliances usage and their energy consumption by monitoring electric plug loads. The mined patterns can be used to reduce the consumption of energy in further studies. The approach is to create a model that contains the characteristics of relations between devices. To create the required models Markov models are used. Markov models can make predictions of the future based on the current state of the model as represented in Figure 2. The figure shows the probability of what the next state will be based on the current state.

[1] Centraal Bureau voor de Statistiek: http://www.cbs.nl
[2] EIA: https://www.eia.gov/

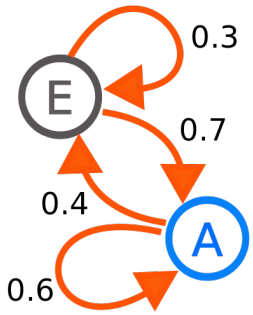

Figure 2: Example of a Markov model[3]

The Markov models will be used to show relations between devices in the form of sequences, where a sequence is the usage of a device after another device. The created models will be used for predicting whether a device will be used or not in the near future, the energy that will be consumed by the device and whether or not it is possible to predict when a device will be used without taking the state of the device itself into consideration.

1.1 RESEARCH QUESTION

The main focus of the current research focuses on energy consumption for an entire building over a longer time period instead of real-time for single appliances. This thesis tries to mine patterns for other studies to use to reduce energy consumption in households.

This leads to the major research question: *'How can patterns be mined for household appliances by monitoring electric plug loads?'*. To answer this question the research question is split up in multiple sub-questions which should help us answer the main question. The sub-questions are:

1. Can a device state be forecasted by monitoring plug loads?
2. Can a device his upcoming energy consumption be predicted by monitoring plug loads?
3. Can a device his state still be forecasted by monitoring plug loads once you stop monitoring the device itself?

Each of the sub-questions will be tried to answered in an experiment. By combining the results of those experiments it should be possible to answer the major research question.

1.2 DOCUMENT STRUCTURE

The remainder of this document is as follows. Chapter 2 describes the model and the algorithms needed to make predictions. In Chapter 3 the data sets used for our experiments are listed and described. After the data sets it is time for the experiments. We start with an experiment to predict device states in Chapter 4 followed by the experiment to forecast a device his energy consumption in Chapter 5. The experiments conclude with a fault tolerance experiment in Chapter 6. After the experiments the related work for state prediction, energy consumption prediction and fault tolerance in sensor networks is reported in Chapter 7. Finally in Chapter 8 the research question is answered and future work is presented.

[3] Taken from: https://upload.wikimedia.org/wikipedia/commons/thumb/2/2b/Markovkate_01.svg/563px-Markovkate_01.svg.png

2

MODEL DESIGN AND ALGORITHMS

Models are an important feature for revealing hidden patterns in big data sets. For this thesis we need a model which characterises relations between devices in a smart building. With the use of this model it should be possible to make predictions on situations in the near future. Following is a definition of this model together with three algorithms which will be used to conduct experiments.

2.1 MODEL

The model uses a technique used in recommender systems to mine frequent sequences. A common approach to mine sequences is the use of Markov models. A Markov model allows for forecasting events based on the current state of the model. As an example if the current state of the model is x then the next state probably will be y. For this thesis a Markov model is used which is based on a solution proposed by Deshpande and Karypis in 2004 where they have used Markov models to predict Web-page access [18]. It contains two phases: (1) the training phase, (2) the prediction phase. During the training phase old, stored, transactions are used to train the model. The results are stored in a SQLite database so they can be used later for real-time predictions.

After training the model contains the probabilities that a device will be used in the next m time-units given a sequence.

Lets take for example a building with the following devices monitored: a fridge, a printer, a coffee machine, a microwave and a screen. Consider the following sequence: $Fridge \Rightarrow Microwave \Rightarrow Printer$. This suggests that after the fridge has been used the microwave has been used and finally the printer has been used. Based on this information the model is able to make a prediction which devices will be used in the next m time-units.

An example of a model is shown in Figure 3 which shows a possible model for two devices and a maximum sequence length of 3. The lines between the nodes tell how often a sequence occurs while the lines to the leafs give the probability that the given device will be used in the next m time-units given the sequence leading to the leaf. In this figure the red nodes are not supported sequences which will be pruned to reduce the complexity of the model as will be described later.

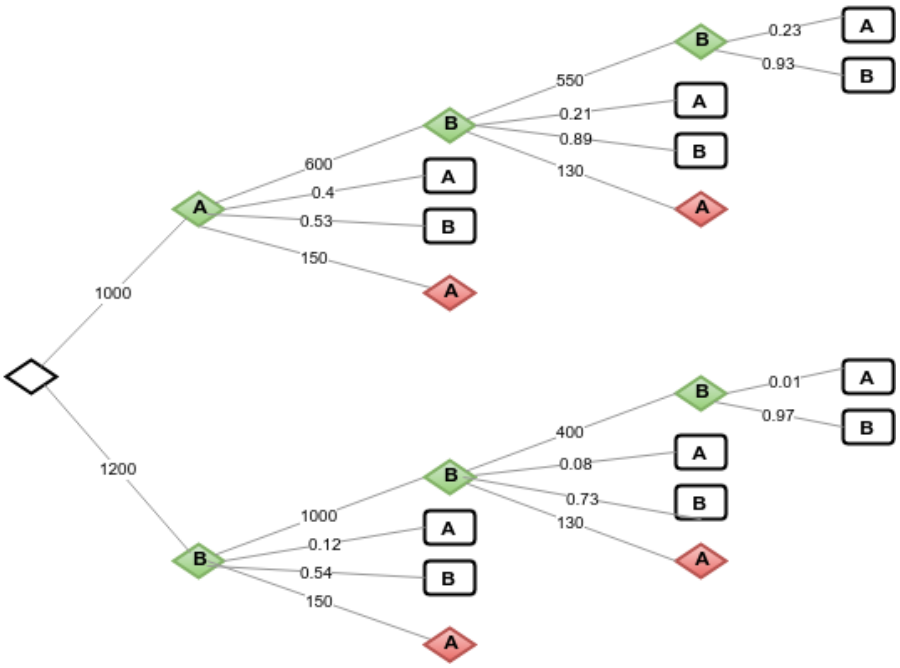

Figure 3: Example of a model with two devices

In the following sections the preliminaries for the algorithm to create this model will be reported followed by the actual algorithm.

2.1.1 *Preliminaries*

Markov models have been widely used for predicting the next action of a user based on his last actions. As in the situation of prediction device states there are some differences compared to discrete user actions some modifications have been applied. First-order Markov models tend to be not successful in predicting further actions and therefore higher order Markov models, Markov models which include more historical states, are needed. Implementing higher order Markov models has a couple of limitations with high state-space complexity, reduced coverage and sometimes worse prediction accuracy [18]. To limit the high state-space complexity a technique is used to prune the set of possible states. This pruned set is later used to predict the actual device states. The technique used is an adapted version of support-pruned Markov models (SPMM) [18]. Support-pruned Markov models suggest that sequences with a support less than a defined threshold are pruned from the set. The solution suggested in [18] uses an absolute threshold, the frequency threshold. In this thesis a relative support threshold is used as proposed in [7]. To overcome the limitation of reduced coverage a technique is used to train multiple order Markov models and use all of them in the prediction. This is called the All-K^{th}-Order Markov model [41].

Following is a list of definitions needed for the algorithm.

DEFINITION DEVICES The algorithm takes a set of devices $\mathcal{V} = \{v_1, ..., v_{|V|}\}$. At any given time t those devices have a state s which can either be on ($s = 1$) or off ($s = 0$).

DEFINITION HISTORICAL STATES Devices their historical states is defined as $\mathcal{S}$ where $\mathcal{S}_{v_i} = \{< t_1, s_1 >, ..., < t_n, s_n >\}$ contains the historical states of device v_i.

DEFINITION TRANSACTION A transaction τ is defined as a set of devices which state $s = 1$ on time t. Giving $\tau(t) = \{v_i | \forall i : \mathcal{S}_{v_i}(t) = 1\}$.

DEFINITION TRANSACTION SET Following the definition of a transaction τ a set of transactions $\mathcal{T}$ is an ordered set of transactions such that $\forall i, j : i < j, \mathcal{T}_i < \mathcal{T}_j$, meaning that all transaction in the set $\mathcal{T}$ are ordered by time in ascending order.

DEFINITION SEQUENCE Given a set of transactions $\mathcal{T}$ a sequence is the transition between the transactions $\tau \in \mathcal{T}$. Consider $\tau_1 = \{A\}$, $\tau_2 = \{B, C\}$ and $\tau_3 = \{A\}$ than τ_1 transits to B and C from τ_2 and to A from τ_3. Using this approach the set $\mathcal{T}$ has the following sequences: $\{ABA\}, \{ACA\}, \{AB\}, \{AC\}, \{A\}, \{B\}$, $\{C\}$, $\{AA\}$, $\{BA\}$ and $\{CA\}$.

DEFINITION SUPPORTED SEQUENCE Using the definition of a sequence as defined above a sequence is supported if its support is above r, the minimal support threshold. The support of a sequence is defined as $Supp(seq) = \frac{w(seq)}{w(p)}$, where w is the number of times a sequence occurs in the transactions set $\mathcal{T}$ and p is the parent sequence of seq such that $p = \{seq_i | \forall i : i < |seq|\}$.

DEFINITION MAXIMUM TIMESPAN The maximum timespan between the first element in $\mathcal{T}$ and the last element ($\mathcal{T}_0 - \mathcal{T}_{|T|}$) is defined as m. The value m can be any arbitrary value but in this thesis m is always measured in minutes.

MAXIMUM SEQUENCE LENGTH To further limit the amount of possible sequences, next to the restriction rate, a maximum length of a sequence is defined. Let $|seq| \leq k$ where k is the maximum length of a sequence. Combining the definition of m it is clear that $k \leq m$.

2.1.2 *Training algorithm*

The algorithm has to be trained by providing historical transactions. From those transactions the probability of each, supported, sequence is computed. The mined probabilities can later be used to make real-time predictions of device usages. The training algorithm (shown in Algorithm 1) takes the set of historical transactions, a restriction rate, a maximum sequence length and the maximum timespan. From this data set the training algorithm will mine for each sequence in the data set the probability that a device will be used in the next m time-units if a sequence occurs.

To limit the number of sequences to train the restriction rate is used in Algorithm 2 to remove non supported sequences from the training.

The function $findOccurrences$ (Algorithm 3) returns all the devices that are on in all the transactions such that $\tau.time < new.time < \tau.time + m$.

Algorithm 1 Training

Require: set of historical transactions $\mathcal{T}$
Require: restriction rate r
Require: Maximum sequence length k
Require: Maximum timespan m

$c \leftarrow \emptyset$
for $i \leftarrow 1$ to $|\mathcal{T}|$ **do** ▷ create first order chain
  $\tau \leftarrow \mathcal{T}[i]$
  for each $v \in \tau$ **do**
    $c \leftarrow$ add occurrence (v, i, i)
  end for
end for
for $i \leftarrow 1$ to k **do** ▷ create remaining orders
  $\mathcal{Q} \leftarrow$ supportedSequencesForOrder(c, i)
  for each $q \in \mathcal{Q}$ **do** ▷ q is a node from c
    for each $o \in q.occurrences$ **do**
      $rootIndex \leftarrow o$.rootIndex
      $transactionIndex \leftarrow o$.index
      $\mathcal{O} \leftarrow$ findNextOccurrences$(rootIndex, transactionIndex + 1)$
      for each $o \in \mathcal{O}$ **do**
        $q \leftarrow$ add occurrence $(o.v, o.index, rootIndex)$
      end for
      $\mathcal{F} \leftarrow$ findFutureOccurrences$(rootIndex)$
      for each $f \in \mathcal{F}$ **do**
        $q \leftarrow$ add future occurrence $(f.v)$
      end for
    end for
  end for
end for
return c

Algorithm 2 Filtering supported sequences

Function supportedSequencesForOrder(c - chain, r - restriction rate, l - sequence length)
$\mathcal{Q} \leftarrow \emptyset$
$\mathcal{C} \leftarrow$ getSetOfChildren(c)
for each $c \in \mathcal{C}$ **do**
  if $|c| = l$ and $c.probability > r$ **then**
    $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{c\}$
  else if $l > |c|$ **then**
    $\mathcal{Q} \leftarrow \mathcal{Q} \cup$ supportedSequencesForOrder(c, l)
  end if
end for
return $\mathcal{Q}$

Algorithm 3 Find next occurrences

```
1: Function findOccurrences(root - rootIndex)
2: T_next ← ∅
3: τ_root ← T[rootIndex]
4: i ← rootIndex + 1 ▷ Find all transactions within the maximum timespan
   m
5: while i < |T| do
6:     τ_tmp ← T[i]
7:     if |τ_root.time − τ_tmp.time| < m then
8:         T_next ← T_next ∪ {τ_tmp}
9:     else
10:        break
11:    end if
12:    i ← i + 1
13: end while
14: R ← ∅
15: for each τ ∈ T_next do
16:    for each d ∈ τ do
17:        R ← R ∪ d
18:    end for
19: end for
20: return R
```

2.2 DEVICE STATE PREDICTION

Predicting which devices will be used in the next m time-units is a possible use-case for the before described model. The illustration in Figure 4 shows an example of a time line with $m = 3$. It tells us that the device A is on now at $t = 0$, device C at $t - 1$ and device A and B are on at $t - 2$. Now Algorithm 4 is proposed to predict what to fill in at the question marks by using the model. There has to be noted that there is no difference between $t + 1$, $t + 2$ or $t + 3$. The algorithm only tries to predict whether a device will be used somewhere in those m time-units but it does not tell at which exact point.

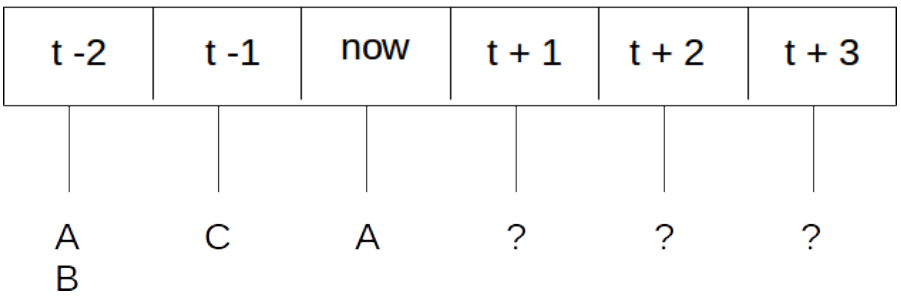

Figure 4: Illustration of a transactions time line

The only data that is available to predict the state of the devices in the next m time-units is the trained database from section 2.1 and the state of the devices at $t = 0$, $t - 1$ and $t - 2$. The algorithm takes a stream of (real time) transactions and keeps the last x transactions such that the time span between $\mathcal{T}[0]$ and $\mathcal{T}[x]$ is less than m time units. Using those cached transactions all possible sequences are retrieved (in the example from Figure 4 this would be $\{A, C, A\}, \{A, C\}, \{A\}, \{B, C, A\}, \{B, C\}, \{B\}, \{A, A\}, \{B, A\}, \{C, A\}$ and $\{C\}$) as $\mathcal{Q}$.

For each device the certainty is then predicted by taking the probability from the before mentioned sequences taken from the trained data set. The certainty $\mathcal{P}$ can be predicted in two ways either by taking the average

Algorithm 4 Device state prediction

Require: Trained data set - db
Require: Transaction stream - $\mathcal{T}_{stream}$
Require: Maximum time units - m
Require: Set of devices - V
1: $\mathcal{T}_{current} \leftarrow \emptyset$
2: **while** $\tau = \mathcal{T}_{stream}.new()$ **do**
3: $\quad\mathcal{T}_{current} \leftarrow \mathcal{T}_{current} \cup \tau$
4: $\quad$**while** $\tau.time - \mathcal{T}_{current}[0].time > m$ **do**
5: $\quad\quad\mathcal{T}_{current} \leftarrow \mathcal{T}_{current} \setminus \mathcal{T}_{current}[0]$
6: $\quad$**end while**
7: $\quad\mathcal{Q} \leftarrow supportedSequences()$
8: $\quad$**for each** $v \in \mathcal{V}$ **do**
9: $\quad\quad certainty \leftarrow$ getCertainty(db, v, $\mathcal{Q}$)
10: $\quad\quad$**if** $certainty > 0.5$ **then**
11: $\quad\quad\quad v_{state} \leftarrow$ on
12: $\quad\quad$**else**
13: $\quad\quad\quad v_{state} \leftarrow$ off
14: $\quad\quad$**end if**
15: $\quad$**end for**
16: **end while**

of the supported sequences or by taking the weighted average. For $\mathcal{P}_{avg}$ the certainty is calculated by taking the sum of all supported sequences probabilities divided by the number of supported sequences. The weighted average method takes the number of times a sequence occurs in the data set into account. Resulting $\mathcal{P}_{wavg}$ is calculated by taking the sum of all supported probabilities $\times$ w and finally divided by the number of supported sequences. Finally if $\mathcal{P}$ is greater than 0.5 the device is predicted to be used in the next m time-units and not used otherwise.

2.3 ENERGY CONSUMPTION PREDICTION

This algorithm continues on the algorithm in section 2.2 by taking the output from that algorithm to take further actions. If a device is predicted to be used (as in line 10 from Algorithm 4) the proposed algorithm tries to predict the amount of Watt it will consume in the next minute.

For this purpose a second Markov model is trained. This model characterises the transitions between power consumptions per device. Each device his power consumption is manually defined in bands and the actually consumption is rounded to the nearest band. From those rounded consumption measurements a Markov chain is created of the last three measurements. The algorithm is listed in Algorithm 5.

As example take device x with as bands $\{0W, 5W, 10W, 30W\}$ and a consumption history of $\{1.2W, 4.3W, 12W, 19W\}$. The consumption history will be rounded to the nearest band resulting in $\{0W, 5W, 10W, 10W\}$. The last three consumptions, excluding the outcome, are $0W, 5W$ and $10W$ so the sequence $0 - 5 - 10$ is trained with an outcome of $10W$. After the model is trained it contains per device for each sequence the most likely next power consumption given in Watt.

Using the outcome of the state prediction algorithm and the model containing the energy consumption transitions it now is possible to predict a device his upcoming energy consumption. The approach is the same as with training the energy consumption model. The observed energy consumptions are rounded to the nearest band and a sequence is created from the last three measurements. As the algorithm is based on the result of the state prediction algorithm it takes that result as input as shown in Algorithm 6. When the state prediction algorithm predicts that the device will be used the energy

Algorithm 5 Energy prediction model training

Require: Device consumption measurements history - $\mathcal{H}$
Require: Set of consumption bands - $\mathcal{B}$
1: $recentHistory \leftarrow \emptyset$
2: $\mathcal{S} \leftarrow \emptyset$ ▷ HashMap(String,Int)
3: **for each** $h \in \mathcal{H}$ **do**
4: $roundedConsumption \leftarrow -1$
5: $tmp \leftarrow 0$
6: **while** $tmp < |\mathcal{B}|$ **do**
7: **if** $h > \mathcal{B}[tmp++]$ and $h < \mathcal{B}[tmp]$ **then**
8: **if** $h - \mathcal{B}[tmp-1] < \mathcal{B}[tmp] - h$ **then**
9: $roundedConsumption \leftarrow \mathcal{B}[tmp-1]$
10: **else**
11: $roundedConsumption \leftarrow \mathcal{B}[tmp]$
12: **end if**
13: **end if**
14: **end while**
15: **if** $|recentHistory| > 3$ **then**
16: $recentHistory \leftarrow recentHistory \setminus recentHistory[0]$
17: **end if**
18: $recentHistory \leftarrow recentHistory \cup roundedConsumption$
19: $s \leftarrow recentHistory.concat('-')$
20: **if** $\mathcal{S} \cap s == \emptyset$ **then**
21: $\mathcal{S} \leftarrow \mathcal{S} \cup s$
22: **end if**
23: $\mathcal{S}[s] \leftarrow \mathcal{S}[s] + 1$
24: **end for**
25: **return** $\mathcal{S}$

consumption model is used to retrieve the expected energy consumption. This is done by selecting the most occurring energy consumption from the model based on the current sequence. When the current sequence is not present in the model the algorithm will use the last as a fail-safe. If the state prediction algorithm decided that the device will not be used the lowest band is selected as most likely energy consumption.

Algorithm 6 Predict energy consumption

Require: Energy model - $\mathcal{E}$
Require: Current consumption sequence - s
Require: State prediction outcome - $state$
Require: Set of bands - $\mathcal{B}$
Require: Last consumption - l
1: **if** $state == off$ **then**
2: **return** $\mathcal{B}[0]$ ▷ Device predicted off. Return smallest band.
3: **end if**
4: **if** $\mathcal{E} \cap s == \emptyset$ **then**
5: **return** l ▷ Unknown sequence. Return last consumption.
6: **end if**
7: **return** $max(\mathcal{E}[s])$ ▷ Return most occurred next consumption

2.4 FAULT TOLERANCE PREDICTION

This algorithm is the same as in section 2.2 with a small modification. To test if the created model is still capable of predicting device states when the device sensor is removed we ignore sequences containing the device to predict.

The final algorithm is listed in Algorithm 7 where the difference between Algorithm 4 is in line 8. In this line the active supported sequences are retrieved, just as in the normal state prediction, but all sequences which contain the device to predict are ignored. This simulates the same behaviour as when a sensor failed or is removed.

Algorithm 7 Device fault tolerance state prediction

Require: Trained data set - db
Require: Transaction stream - $\mathcal{T}_{stream}$
Require: Maximum time units - m
Require: Set of devices - V
1: $\mathcal{T}_{current} \leftarrow \emptyset$
2: **while** $\tau = \mathcal{T}_{stream}.new()$ **do**
3: $\quad\mathcal{T}_{current} \leftarrow \mathcal{T}_{current} \cup \tau$
4: $\quad$**while** $\tau.time - \mathcal{T}_{current}[0].time > m$ **do**
5: $\quad\quad\mathcal{T}_{current} \leftarrow \mathcal{T}_{current} \setminus \mathcal{T}_{current}[0]$
6: $\quad$**end while**
7: $\quad$**for each** $v \in \mathcal{V}$ **do**
8: $\quad\quad\mathcal{Q} \leftarrow supportedSequences() \setminus sequencesWith(v)$
9: $\quad\quad certainty \leftarrow$ getCertainty(db, v, $\mathcal{Q}$)
10: $\quad\quad$**if** $certainty > 0.5$ **then**
11: $\quad\quad\quad v_{state} \leftarrow$ on
12: $\quad\quad$**else**
13: $\quad\quad\quad v_{state} \leftarrow$ off
14: $\quad\quad$**end if**
15: $\quad$**end for**
16: **end while**

3 DATA SETS

Data sets are necessary to create and test the model. A significant amount of power consumptions measurements is needed in order to train the model and make predictions based on the model. For the purpose of this thesis three data sets have been obtained where two data sets are from the same experiment.

3.1 RUG DATA SET

The RUG data set is mined from an experiment in the Bernoulliborg building of the University of Groningen. The data set has five devices: a boiler, a coffee maker, a printer, a microwave and a TV-screen. Those devices are all part of a coffee corner on the fifth floor of the mentioned building. A layout of the coffee corner is available in Figure 5. Power consumption measurements have been recorded using Plugwise adapters [3] with 6 measurements per minute from 03-02-2016 till 15-08-2016.

As a measurement per ten seconds is too often for this thesis the data set has been pruned to one measurement per minute. To achieve this reduction the maximum energy consumption per minute is used. This results in a data set of 217.452 transactions. There are 279.360 minutes in the given period which tells us that there are moments where no data is available and the data set has a coverage of 77,84%.

The data set has been split into two parts to get a training and a validation set. The training set ranges from 03-02-2016 till 09-06-2016 and has 131.386 transactions. The remaining data, 86.066 transactions, are used as the validation data set and contains measurements from 09-06-2016 till 15-08-2016.

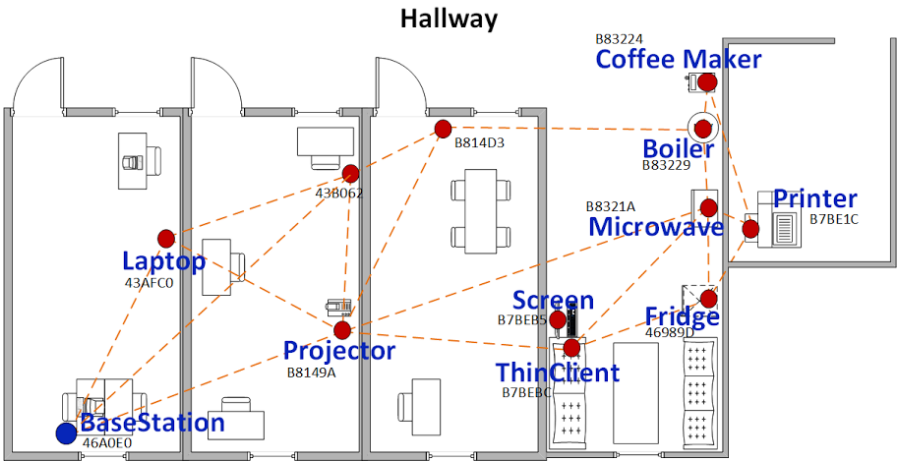

Figure 5: Floor plan of the coffee corner in the Bernoulliborg building

3.2 ELECTRONIC CONSUMPTION OCCUPANCY DATA SET

The Electronics Consumption Occupancy (ECO) data set [33, 13] is a "comprehensive data set for non-intrusive load monitoring and occupancy detection research" created by the Swiss Federal Institute of Technology in Zurich. The data was collected during six months for the purpose of the project Smart

Meter Services [1]. The ECO data set contains six Swiss households from which two were picked to use in this thesis.

3.2.1 *Household 1*

The first household, from now on named household 1, monitors the power consumption of four devices. It has a fridge, a washing machine, a dryer and a freezer. It contains $1Hz$ consumption data from 01-06-2012 till 23-01-2013 with a coverage of 99.64%. The data set is, as before, again reduced to one measurement per minute. This data set has the limitation that the data from the first 29 days (till 29-06-2012) is rounded to $10W$ precision. The data set is split in a training and validation set where the training set contains 211.619 transactions and the validation set has 120.960 transactions.

3.2.2 *Household 2*

Household 2 is the biggest data set used. The data set contains $1Hz$ consumption measurements data of twelve devices in the period from 01-06-2012 to 31-01-2013 and has a coverage of 98.58%. While trying to train the model on this data set it came to light that with twelve devices there are simply too many possible sequences which makes it impossible to compute. It has to be said that the number of trainable devices matters on the density of the data set. If devices are not often used it is easier to train the model as sequences are more likely to not hit the restriction rate and thus be pruned.

Finally seven devices from the data set have been selected to be used. This are a dishwasher, an air exhaust, a fridge, a freezer, a dimmable lamp, a TV and as last a stereo. Again the data set is reduced to one minute measurements and split to have a training and validation set. The training set is from 01-06-2012 till 30-09-2012 and contains 175.679 transactions while the validation set is from 01-10-2012 till 31-01-2013 and has 169.860 transactions.

4 EXPERIMENT ONE: STATE PREDICTION

Predicting device states in the next m time-units can potentially be used to save energy. If a device is predicted to not be used it can be put into standby mode which can save up to 10% of the energy consumption [51] or even disabled completely. In this experiment we use the model and the method described in Section 2.2 to try to predict device usage. This approach is conducted for all the data sets available in Chapter 3.

For this experiment the parameters of the model have to be tweaked first to gain the best possible results. After the parameters have been tweaked the results will be discussed to finish this chapter with a small conclusion.

4.1 PARAMETER TWEAKING

As it is impossible to test and validate all possible parameter combinations each parameter is isolated to test the impact of a single parameter. This is done by changing one parameter while keeping the other parameters constant. The parameter tweaking is only performed on the RUG data set as no, significant, differences are expected to occur for other data sets.

4.1.1 *Minimal support threshold r*

The graph in Figure 6 shows the effect of the minimal support threshold. This is done by comparing the number of supported sequences per threshold. The more supported sequences the higher the complexity of the model.

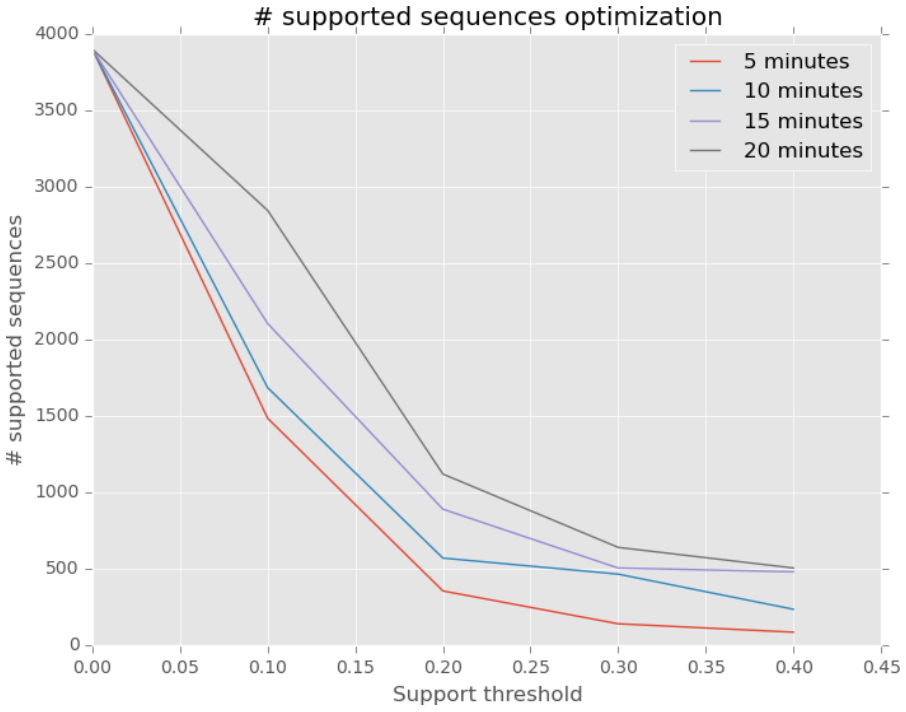

Figure 6: Effect of the minimal support threshold on # supported sequences. Mined with $|\mathcal{T}| = 131386$, $|V| = 5$ on the RUG data set

As can be seen in Figure 6 the support threshold has an huge impact on the number of supported sequences. With only a small threshold (< 0.3) the number of supported sequences already drops significantly. The same

test has been done but now including devices in the trained sequences that are off. As a result the number of supported sequences grew drastically. At $k = 5, m = 5$ and $r = 0.2$ there were 15.290 supported sequences, compare that to 355 supported sequences when only including devices that are on and you can understand that this is an undesired approach.

Now it is interesting to compare the drop of number of supported sequences to the accuracy of predictions. As can be seen in Figure 7 the difference in accuracy is minimal for all $r < 0.4$. The difference between $r = 0.0$ and $r = 0.2$ is only 0.16%. A possible explanation for this observation is that sequences with a low support (e.g 0.1) are so unlikely to occur in predictions that their probabilities have barely any influence on the final predictions.

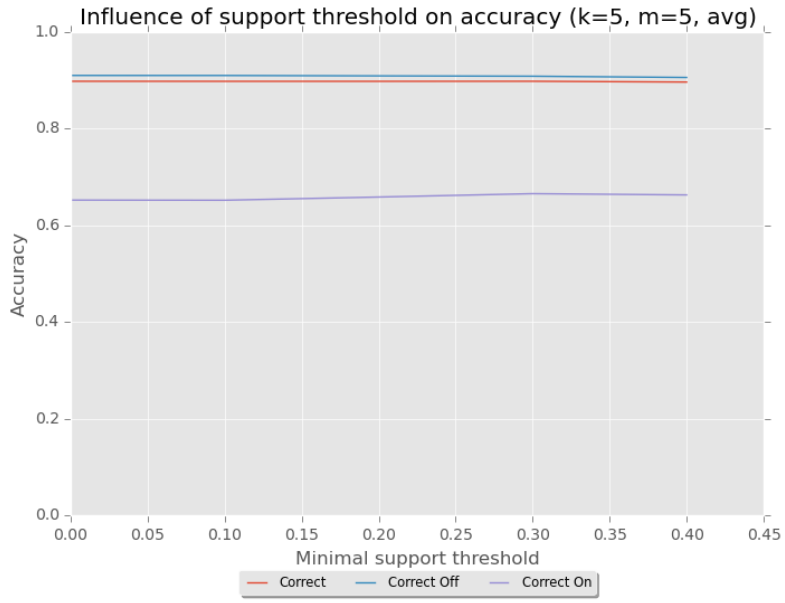

Figure 7: Effect of support sequence on accuracy. ($r = 0.3$, $m = 10$)

4.1.2 *Sequence length k*

To retrieve the impact of the maximum sequence length on the accuracy a test run is made. This test is made with a timeslot of $m = 10$ minutes, minimal support threshold $r = 0.3$ and the average predictor method. The results, as shown in Figure 8, show a pretty stable accuracy for all $k \geq 2$, $k \leq 10$. An interesting observation is the drop in the accuracy for predicting that a device will be used. At $k = 2$ the accuracy for predicting that a device will be used is around 58% while at $k = 3$ it is only 40%. This stabilizes at $k \geq 5$ with an accuracy of 45%.

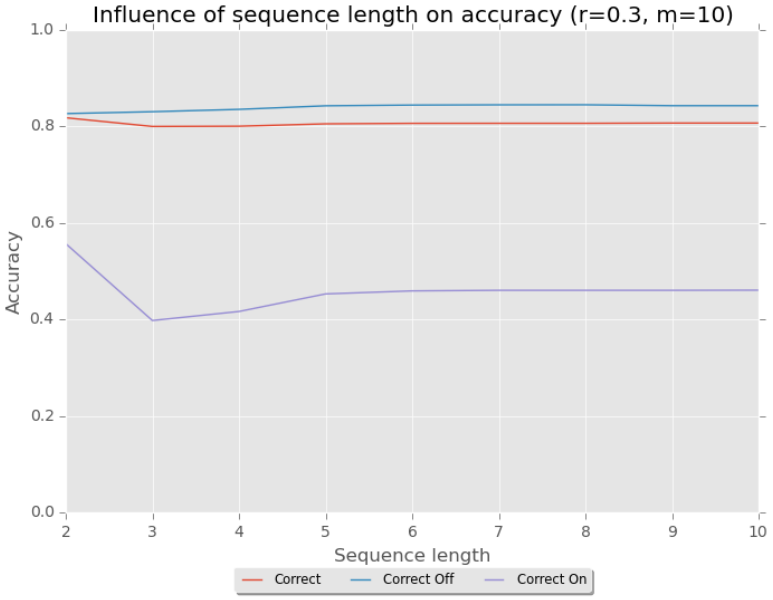

Figure 8: Effect of maximum sequence length ($m = 10$, $r = 0.3$, average)

4.1.3 *Time-unit lenth m*

In Figure 9 the results of tweaking the time length variable (m) is shown. As the graph shows a weird, unexplainable, curve additional information is added to the graph. Instead of only showing the correct percentage also the correct off, correct on percentages are shown. As the accuracy still is hard to explain the percentages of real on/off information is added. Overall the graph shows that the accuracy is best for either a small value as $m = 5$ or a big value $m \geq 20$. With an average time length ($m = 10$) the accuracy, especially for predicting an appliance on, is worse. This can be explained as with a high m the change of a device being used is a lot higher and thus is it easier to predict this. At the same time with a low m the change of being an appliance used is small and is it unlikely that it actually is used.

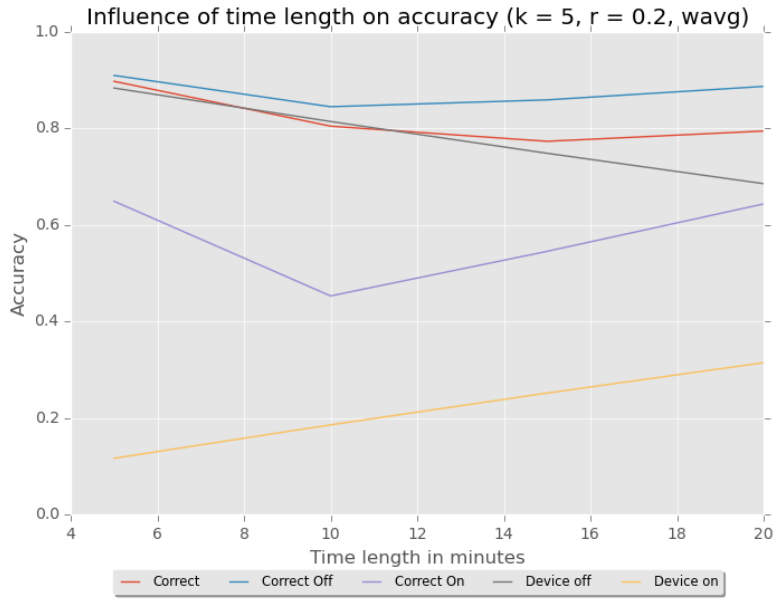

Figure 9: Effect of time length ($k = 5$, $r = 0.2$, average)

4.1.4 *Predictor method*

There are two algorithms to calculate the probability that a device will be used in the next m minutes. Method one takes the average of all supported sequences and method two takes the weighted average of the supported sequences. As can be seen in Figure 10 there is barely any difference in the accuracy of the predictions between the two methods. For simplicity we stick to the weighted average method for the rest of this thesis unless mentioned otherwise.

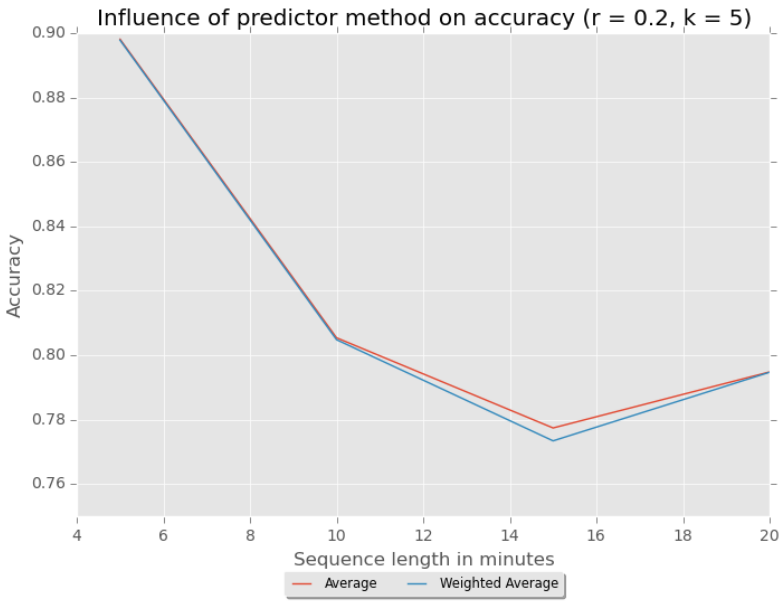

Figure 10: Effect of predictor method ($r = 0.2, k = 5$)

4.1.5 *Optimal configuration*

Following the results from the parameters tweaking a, near, optimal configuration can be derived. The results suggest an optimal configuration of $m = 5$, $r = 0.2$, $k = 5$ and as predictor method *wavg*. This configuration will be used to validate the results on the available data sets.

4.2 RESULTS

In Table 1 the results of the data sets are compared. It shows comparable results for each data set with an accuracy high in the 80% while scoring better for predicting that a device will not be used, all above 90%, than that a device will be on, varying between 65% and 77%. A possible reason for the difference in accuracy between predicting on and off is the amount of time devices are on; the more time devices are on the higher the accuracy of predicting a device correctly on.

Statistic	RUG	ECO1	ECO2
Correct	0,8978	0,8892	0,8786
Correct off	0,91	0,9458	0,9628
Correct on	0,6492	0,7707	0,7179
Wrong	0,1023	0,1109	0,1215
Device off	0,8838	0,7145	0,7289
Device on	0,1163	0,2856	0,2712
On coverage	0,2617	0,8713	0,9098
Off coverage	0,9815	0,8964	0,8671

Table 1: Comparison of the different data sets

In the following sections each data set is individually validated by describing the data set and reporting the same values as in Table 1 but than device specific. This should give insight in what kind of devices are better to predict and what kind of device might cause problems.

4.2.1 *RUG data set*

In Table 2 the device specific results are listed. In some cases this shows a really high accuracy, especially for the screen, microwave and boiler. However when looking at the on coverage the results are unsatisfying. The microwave is the only device whose on coverage is higher than 50% while for the boiler and coffee maker this is even 0%. The histogram in Figure 11 shows that the model is in more of 60% of the times sure that a device will not be used. This matches with the information from Table 2 which shows that most devices are barely ever on. A more detailed validation of the model on this data set is available in section A.1. In this appendix the devices are validated individually.

Statistic	Screen	Microwave	Boiler	Coffee maker	Printer
Correct	0,9998	0,967	0,9691	0,8077	0,7453
Correct off	0,9998	0,9734	0,9691	0,8077	0,79
Correct on	0,9334	0,9084	0	0	0,4588
Wrong	0,0003	0,0331	0,031	0,1924	0,2548
Device off	0,9996	0,8861	0,9691	0,8077	0,7564
Device on	0,0005	0,114	0,031	0,1924	0,2437
On coverage	0,3889	0,7895	0	0	0,2544
Off coverage	1	0,9898	1	1	0,9034

Table 2: Validation of state predictions on the RUG data set

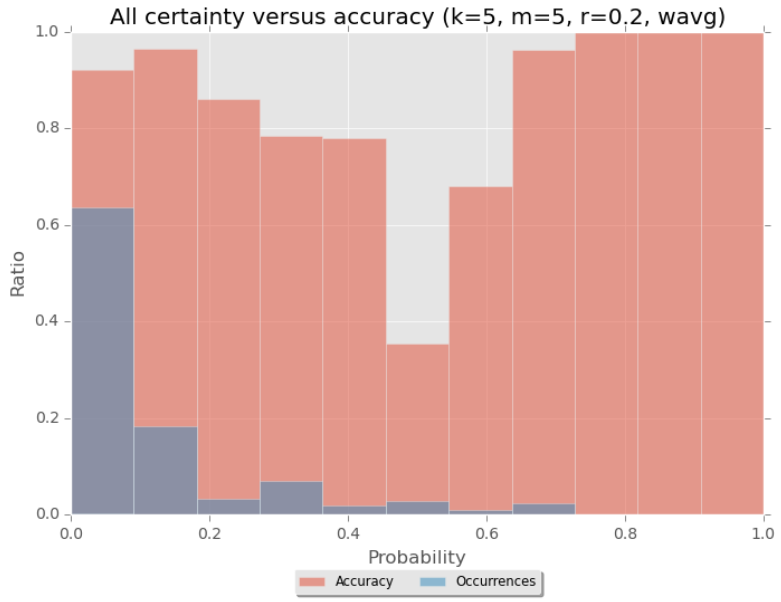

Figure 11: Certainty of state predictions on the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

4.2.2 *ECO data set household 1*

Table 3 reports decent scores for this data set. The device with the worst accuracy, the freezer, still has an accuracy of almost 72% while the dryer and the washing machine have an accuracy of 99% and 97% respectively. For both devices this is achieved by the off coverage which is near 100% but the on coverage is still 82% and 74%.

Statistic	Fridge	Dryer	Washing machine	Freezer
Correct	0,8764	0,9922	0,971	0,7173
Correct off	0,9082	0,9928	0,9749	0,7198
Correct on	0,8303	0,9738	0,9209	0,7167
Wrong	0,1237	0,0079	0,0291	0,2828
Device off	0,6063	0,9606	0,9104	0,3806
Device on	0,3938	0,0395	0,0897	0,6195
On coverage	0,8621	0,8238	0,7397	0,8993
Off coverage	0,8856	0,9991	0,9938	0,4213

Table 3: Validation of state prediction on household 1 of the ECO data set

The certainty plot in Figure 12 shows that the model is mostly certain on this data set. Near 80% of all the predictions are made with a certainty $\leq$ than 0.2 or a certainty $\geq$ 0.8. The graph shows an unexplainable drop in accuracy for predictions with a certainty of 0.7. The accuracy with a certainty of 0.6 reaches an accuracy around 90% while for 0.7 the accuracy is only around 30%. The expected result would be that the predictions with a certainty of 0.7 gain a higher precision as the model is more confident in those predictions. To be able to explain this drop it is required to look at the cause of this effect. This detailed information is available in section A.2.

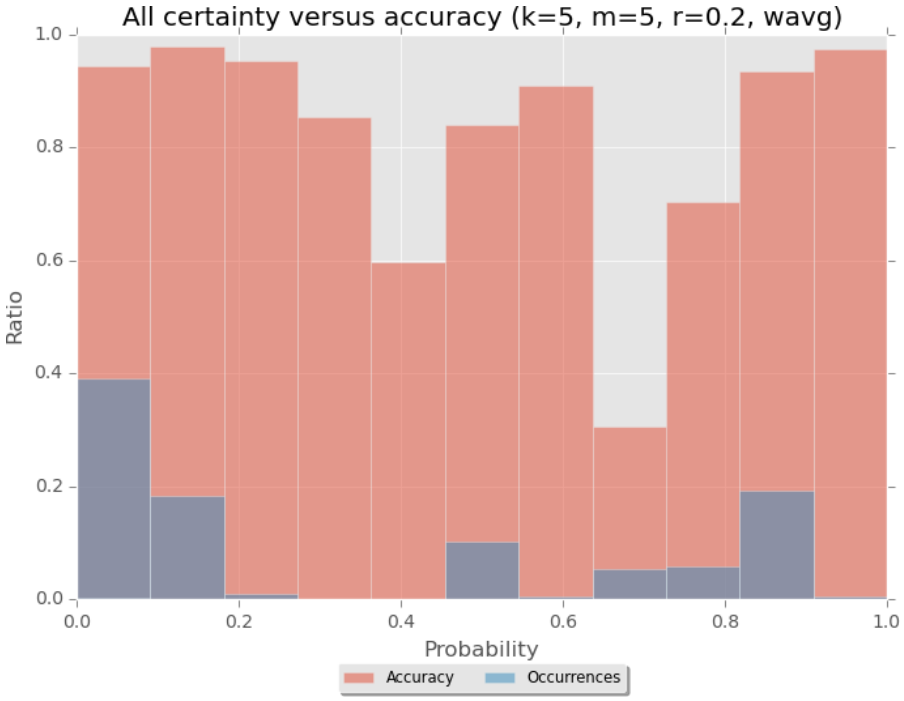

Figure 12: Certainty of state predictions on household 1 ($r = 0.2, k = 5, m = 5, wavg$)

4.2.3 *ECO data set household 2*

The results for this data set are listed in Table 4. As mentioned before in Chapter 3 this data set is more complex than the other data sets. For this reason this data set uses a minimal support threshold of 0.4 to limit the number of supported sequences. The results show a high accuracy for the dish washer, air exhaust, lamp, TV and stereo and a drop in accuracy for the fridge and freezer. The fridge and the freezer especially seems to have a problem detecting that they will not be used as the off coverage is only 52% for the fridge and even 30% for the freezer. The air exhaust, on the other hand, has a problem with predicting that it will be used. A likely reason for this is that it is only on during 1% of the time. Although this sounds feasible the dishwasher is still capable of achieving an on coverage of 60% while the dishwasher is used 1.5% of the time.

Statistic	Dishwasher	Air exhaust	Fridge	Freezer	Lamp	TV	Stereo
Correct	0,9929	0,9898	0,6662	0,6586	0,9852	0,8851	0,9728
Correct off	0,9935	0,9903	0,7733	0,6948	0,9854	0,9962	0,9914
Correct on	0,9364	0,7716	0,6053	0,651	0,9846	0,7024	0,943
Wrong	0,0072	0,0103	0,3339	0,3415	0,0149	0,115	0,0273
Device off	0,9842	0,9886	0,532	0,409	0,8247	0,732	0,6321
Device on	0,0159	0,0115	0,4681	0,5911	0,1754	0,2681	0,368
On coverage	0,5908	0,1475	0,8244	0,9106	0,9303	0,991	0,9855
Off coverage	0,9994	0,9995	0,527	0,2943	0,9969	0,8463	0,9653

Table 4: Validation of state prediction on household 2 of the ECO data set

The certainty of the model for this data set is shown in Figure 13 and shows that half of the time the model is certain that a device absolutely will not be used. The remaining predictions show a, near, even distribution over the remaining possible certainties. The accuracy of the predictions is a bit lower with a certainty of 0.7 compared to 0.6 this is a drop that probably can be explained with the detailed validation in section A.3.

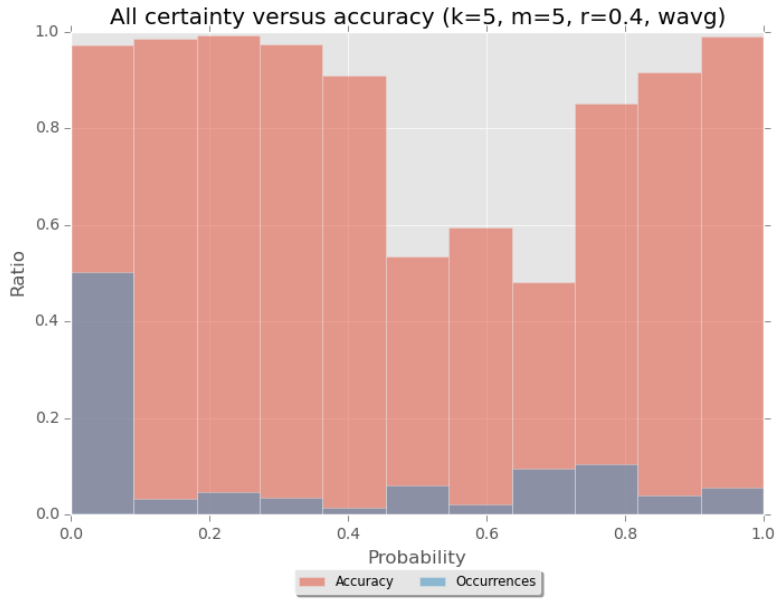

Figure 13: Certainty of state predictions on household 2 ($r = 0.4, k = 5, m = 5, wavg$)

4.3 CONCLUSION

The results from validating the model indicate that the proposed algorithm is suitable for predicting device states in the near future. After testing the algorithm on three different data sets containing a total of 15 devices the accuracy for all data sets is on average high in the 80%. The results tell that not all kind of devices are as easy to predict as other devices. Some devices are barely used and therefore don't show up often in the trained sequences, for example a boiler. Meanwhile other devices run on an automatic schedule instead of being triggered by an user. This is the case for example a fridge and a freezer. This fact makes those devices hard to predict with Markov chains and require a different approach. The model brought to light some hidden patterns between devices which you would not think of such as a relation between a microwave and a printer.

Overall the model predicts more accurate when a device will not be used than that it will be used. The first is, on average, around 70% accurate while the latter has an accuracy of over 90% for all data sets. A reasoning for this is that devices are most of the time not used (around 70% of the time) and thus more likely to guess right.

5

EXPERIMENT TWO: ENERGY CONSUMPTION PREDICTION

By predicting the total energy consumption a user can be warned if he is about to consume too much energy. Imagine living of the grid and you are only getting power from sonar panels and/or windmills. In those situations you have, varying, limited electricity resources. By predicting the upcoming energy consumption it can be potentially foreseen if the energy consumption will be bigger than the produced energy and actions can be taken to prevent power shortages.

In this experiment an attempt is made to predict the energy consumption for each available data set and to compare it with the actual consumption. The technique to predict the energy consumption is proposed in Section 2.3. For comparing the results we take a look at small slices of the energy consumption to look for patterns and use the Mean Absolute Error (MAE) to give a quantitative result. The Mean Absolute Error is a measurement which gives an indication of how close predictions are to the actual observed value. The MAE is measured by the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| \tag{1}$$

where f_i is the prediction at point i and y_i is the actual observation at that moment.

For each device in the data sets a table is provided which shows the MAE, the average energy consumption and a percentage which shows the Mean Absolute Error in percentages with respect to the average energy consumption (MAPE). This percentage is given by the following equation:

$$MAPE = \frac{MAE}{\frac{1}{n}\sum_{i=1}^{n} y_i} * 100\% \tag{2}$$

where y_i is again the observed value at point i.

The lower the MAE and the MAPE the better the results. Following are the results of the energy consumption prediction on the different data sets. For each device there is an individual graph available in Appendix B.

5.1 RUG DATA SET

When looking at the RUG data set plot in Figure 14 it is observable that the algorithm has serious problems in predicting the absolute energy consumption peaks. When the energy consumption is high for a longer time the algorithm is capable of predicting the consumption with a decent accuracy. Because the energy predictions depend on the state predictions the errors in the state predictions have a huge impact on the energy prediction. As the state prediction always fails for the coffee machine and the boiler the energy prediction for those is always close to zero. The coffee machine often uses over 2500W while the boiler uses over 2000W when in use. Those two devices cause already an error of 4500W in the predictions when both are on at the same time and thus cause a huge MAE error. This error rate is visible in the quantitative results in Table 5 where both the coffee maker and the boiler have an MAE which is almost the same as the average consumption. Overall this data set reports a MAE of 90% of the average total consumption.

Device	MAE (W)	Average consumption (W)	MAPE
Screen	0,92	2,1	43,81%
Microwave	28,75	27,67	103,90%
Boiler	35,75	37,22	96,05%
Coffee maker	149,08	152,08	98,03%
Printer	20,58	32,15	64,01%
Overall	227,62	251,23	90,60%

Table 5: Results of energy prediction on the RUG data set

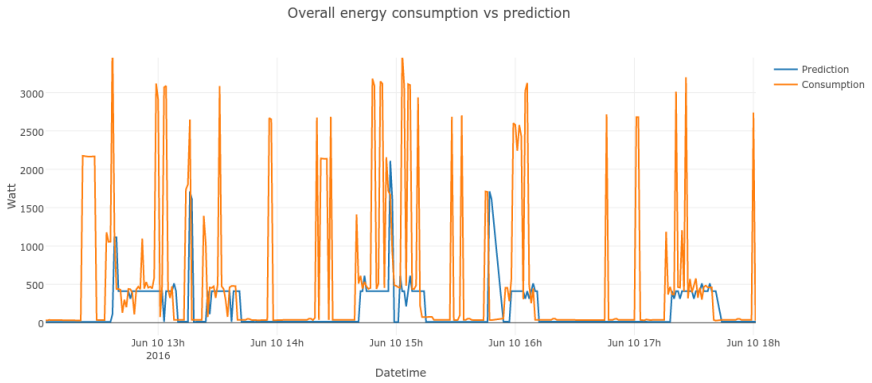

Figure 14: Energy consumption prediction for RUG data set

The results have been achieved with the following bands per device:

- *Screen* $\{0, 100, 200\}$,
- *Microwave* $\{0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200\}$,
- *Printer* $\{0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400\}$.

The boiler and coffee maker have not been configured. The reason for this is that both devices are never predicted on so their bands will not be used a single time.

5.2 ECO DATA SET HOUSEHOLD 1

The result for household 1 of the ECO data set show better results than for the RUG data set. As can be seen in Figure 15 the prediction matches closer with the actual energy consumption than for the RUG data set. The results in Table 6 shows that all devices have a MAPE which is less than 50% of the average consumption with an overall MAPE of 29% of the total average consumption.

Device	MAE (W)	Average consumption (W)	MAPE
Fridge	9,88	25,74	38,38%
Dryer	3,67	29,59	12,40%
Washing machine	19,89	40	49,73%
Freezer	5,23	18,86	27,73%
Overall	33,48	114,15	29,33%

Table 6: Results of energy consumption prediction on household 1

The graph shows that the algorithm can follow the actual consumption when a lot of electricity is consumed. The likely reason is that at the high peaks there is a device running which consumes a lot of electricity for a longer period (e.g the dryer). As the overall state predictions for this data set

are better than for the RUG data set huge errors, as with the boiler and the coffee maker, are less present in this data set.

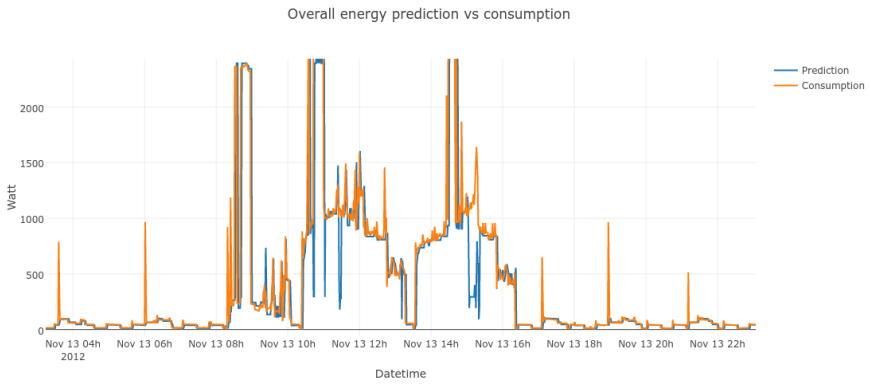

Figure 15: Energy consumption prediction for household 1

To get to this results the following bands were configured per device:

- *Fridge* $\{0, 5, 10, 30, 40, 60, 100, 200, 400, 600, 800, 1000\}$,
- *Dryer* $\{0, 5, 50, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950\}$,
- *Washing machine* $\{0, 100, 200, 300, 400, ..., 2000, 2100, 2200\}$,
- *Freezer* $\{0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80\}$.

5.3 ECO DATA SET HOUSEHOLD 2

Household 2 shows the most stable results as shown in Figure 16. The prediction curve, nearly, matches the actual consumption curve apart from a few small peaks. When looking into the cause of those peaks it is shown that this is completely caused by the fridge. The fridge tends to be off and consumes a lot of electricity in the first minute it comes back on which is shown in Figure 47. The quantitative results listed in Table 7 show some interesting results. It shows that the air exhaust has a MAE of 3*W* but this still causes a 278% error when looking at the MAPE. In this case the MAE of 3*W* gives a better expression than the percentage error as the average consumption is so low that even a small MAE causes a big percentage error. The most interesting results are for the lamp, TV and the stereo with a MAE of only a few Watt and a MAPE of around 10%. The state prediction results from Table 4 show that those devices have an on coverage in the 90%. This tells that there only a few errors passed on from the state prediction to the energy prediction which helps in reducing errors.

Device	MAE (W)	Average consumption (W)	MAPE
Dishwasher	11,16	18,01	61,97%
Air exhaust	3,26	1,17	278,63%
Fridge	33,99	48,04	70,75%
Freezer	33,56	49,74	67,47%
Lamp	3,55	27,09	13,10%
TV	2,55	42,22	6,04%
Stereo	1,91	19,48	9,80%
Overall	79,32	205,71	38,56%

Table 7: Results of energy prediction on household 2

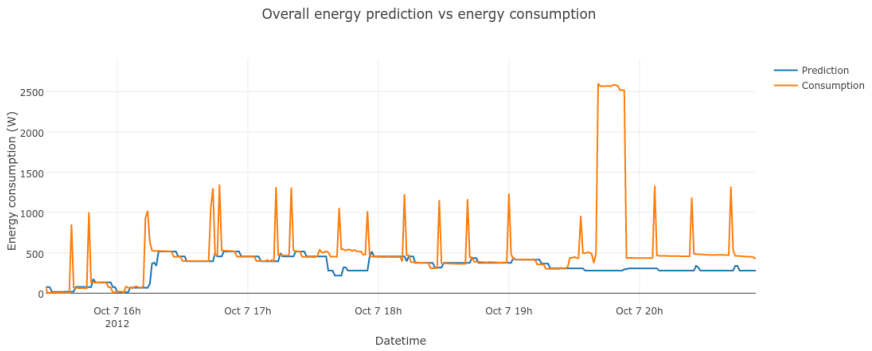

Figure 16: Energy consumption prediction for household 2

The bands for this household were configured as

- *Dishwasher* $\{0, 100, 200, 300, ..., 1500, 1600\}$,
- *Air exhaust* $\{0, 5, 10, 30, 40, 60, 100, 200, 400, 600, 800, 1000\}$,
- *Fridge* $\{0, 5, 10, 30, 40, 60, 100, 200, 400, 600, 800, 1000\}$,
- *Freezer* $\{0, 5, 10, 30, 40, 60, 100, 200, 400, 600, 800, 1000\}$,
- *Lamp* $\{0, 10, 20, 30, ..., 190, 200\}$,
- *TV* $\{0, 2, 5, 140, 150, 160, 170\}$,
- *Stereo* $\{0, 10, 20, 30, ..., 190, 200\}$.

5.4 CONCLUSION

The results for predicting the energy consumption show varying results. Errors from the state prediction are taken over by the energy consumption prediction. Therefore the results of the energy prediction highly depend on the accuracy of the state predictions. When the state predictions quality is low the same will be for the energy predictions as reported for the coffee machine and the boiler from the RUG data set. On the other hand if the state predictions are good as for the TV and stereo from household 2 the results are great.

Overall there are a few conclusions which can be derived from the results.

1. The algorithm performs better on devices which are used often,
2. the algorithm is not good in predicting peaks,
3. the algorithm tends to follow the consumption instead of really forecasting it,
4. the algorithm might perform better with $1Hz$ consumption data but it increases the complexity drastically.

The proposed algorithm is capable of predicting the energy consumption with a decent accuracy but it highly depends on the kind of devices. A device with a regular energy consumption pattern such as a TV or a dimmable lamp will give you high accuracies. At the same time devices as a fridge or a coffee machine tend to be less accurate as they run in a peak pattern which is barely depended on the previous consumptions.

6 EXPERIMENT THREE: FAULT TOLERANCE

Sensors tend to fail due to hardware failure or an error in transmitting the measurements to the central control unit. When a sensor fails the state of the device attached to this sensor will be unknown. In this experiment an attempt is made to predict device states after the sensor of the device has failed. Apart from sensor failure this experiment will give us insight if the number of plugs to be monitored can be limited. If there are devices which show similar results for fault tolerance prediction as with the normal state prediction their sensor possibly could be removed.

For this experiment the approach suggested in Section 2.4 is used to simulate sensor failure. With this approach the actual device states are known so the predictions can be validated. The experiment is performed in the same way as with experiment one (Chapter 4) and thus the results will be reported in the same matter. By performing the results in the same way as with the normal state prediction the results can be compared directly to each other.

The results for this experiment are listed in Table 8 while device specific results are available in Appendix C. The table reports the same data as previous for the state prediction (available in Table 1) with an extra row which tells how many times the prediction is the same as for the normal state prediction. The results show that for the RUG data set the predictions are the same for 95% of the time, about 80% for household 1 and about 90% for household 2.

Statistic	RUG	Household 1	Household 2
Correct	0,8838	0,7402	0,8102
Correct off	0,8838	0,7596	0,8562
Correct on	0	0,6035	0,667
Wrong	0,1163	0,2599	0,1899
Device off	0,8838	0,7145	0,7289
Device on	0,1163	0,2856	0,2712
On coverage	0	0,2628	0,5986
Off coverage	1	0,931	0,8889
Same as ground truth	0,9516	0,7988	0,8998

Table 8: Fault tolerance comparison of the different data sets

6.1 RUG DATA SET

The RUG data set is not capable of predicting device states when a sensor failed. The results from Table 9 show that for all devices the proposed solution is unable to predict the on-state correctly. This results match with the certainty of the predictions from Figure 17 which show that for 80% of the time the algorithm is certain that the device will not be used at all with an accuracy of 90% and shows that the algorithm predicts each device as off for 100% of the time. The reason for this is that the model contains not a single sequence for any device whose probability that the device will be used is higher than 0.5 without containing the device it self in the sequence.

That there is not a single sequence matching this criteria indicates that there is not a strong relation between any of the devices monitored. A possible reason for the lack of such a relation is the total on-time for the devices, this is often only a few percent of the day which makes it unlikely to develop such a relation.

Statistic	Screen	Microwave	Boiler	Coffee maker	Printer
Correct	0,9996	0,8861	0,9691	0,8077	0,7564
Correct off	0,9996	0,8861	0,9691	0,8077	0,7564
Correct on	0	0	0	0	0
Wrong	0,0005	0,114	0,031	0,1924	0,2437
Device off	0,9996	0,8861	0,9691	0,8077	0,7564
Device on	0,0005	0,114	0,031	0,1924	0,2437
On coverage	0	0	0	0	0
Off coverage	1	1	1	1	1
Same prediction	0,9998	0,8932	1	1	0,8649

Table 9: Fault tolerance results for the rug data set

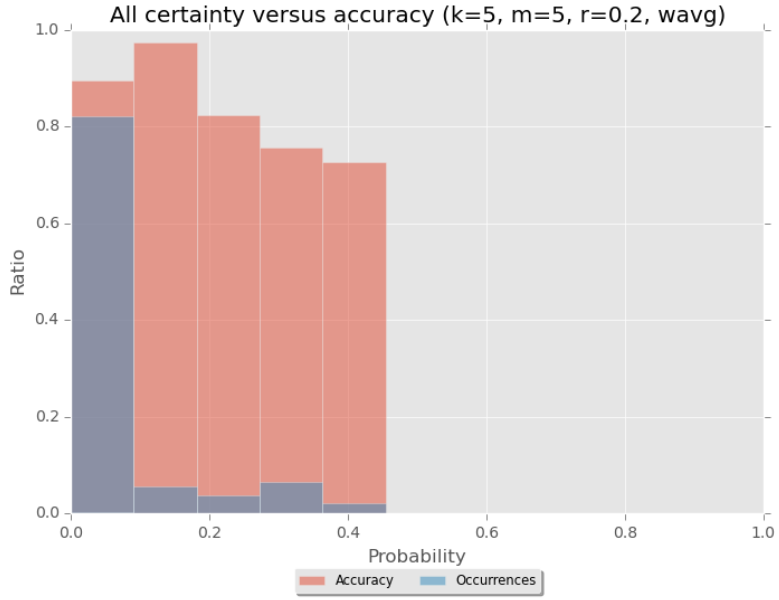

Figure 17: Fault tolerance certainty of predictions on the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

6.2 HOUSEHOLD 1

Household 1 shows slightly better result for fault tolerance prediction than the RUG data set. Table 10 shows that the algorithm is capable of predicting the fridge and freezer to be used while the power measurement of them have been disabled. Although their accuracy is lower (87% versus 60% and 71% versus 49%) it proofs that the algorithm is capable of predicting device states correctly as long as a strong relationships between devices exists. Figure 18 shows that during around 10% of the time the model predicts devices to be used with a certainty of 0.7. With this certainty an accuracy of 60% is achieved which is a bit less as you normally would expect for that certainty.

Statistic	Fridge	Dryer	Washing machine	Freezer
Correct	0,598	0,9606	0,9104	0,4919
Correct off	0,607	0,9606	0,9104	0,3835
Correct on	0,4071	0	0	0,623
Wrong	0,4021	0,0395	0,0897	0,5082
Device off	0,6063	0,9606	0,9104	0,3806
Device on	0,3938	0,0395	0,0897	0,6195
On coverage	0,0463	0	0	0,4552
Off coverage	0,9563	1	1	0,5516
Same prediction	0,636	0,9623	0,9215	0,6754

Table 10: Fault tolerance results for household 1

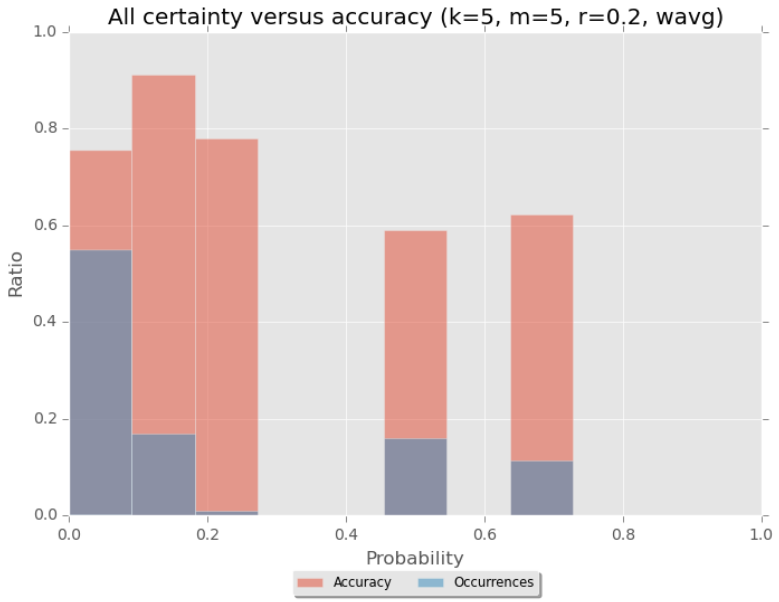

Figure 18: Fault tolerance certainty of predictions on household 1 ($r = 0.2, k = 5, m = 5, wavg$)

6.3 HOUSEHOLD 2

Table 11 reports varying results for the different devices. The real interesting results are achieved for the TV and the stereo. The TV has the same prediction as in the normal state prediction for 99,98% of the time with an overall accuracy of 88,5%. The stereo got even a little bit better results with an overall accuracy of 89% while 92% of the predictions are the same as the ground truth. Overall this data set shows same results as the other data sets. The algorithm is capable of predicting some kind of devices, mostly devices that are used more often, while other devices are impossible to predict. The data in Figure 19 shows that this data set is the only data set where the model sometimes is absolute certain that a device will be used with an accuracy of almost 100%. This tells us that there is one, or multiple device(s) which will be used if another device is used almost all the time.

Statistic	Dishwasher	Air exhaust	Fridge	Freezer	Lamp	TV	Stereo
Correct	0,9842	0,9886	0,544	0,5518	0,8247	0,885	0,8932
Correct off	0,9842	0,9886	0,5619	0,4458	0,8247	0,9959	0,8763
Correct on	0	0	0,5156	0,612	0	0,7025	0,932
Wrong	0,0159	0,0115	0,4561	0,4483	0,1754	0,1151	0,1069
Device off	0,9842	0,9886	0,532	0,409	0,8247	0,732	0,6321
Device on	0,0159	0,0115	0,4681	0,5911	0,1754	0,2681	0,368
On coverage	0	0	0,4248	0,6601	0	0,9905	0,7654
Off coverage	1	1	0,6489	0,3952	1	0,8464	0,9675
Same prediction	0,9901	0,9979	0,7482	0,8107	0,8344	0,9998	0,9177

Table 11: Fault tolerance results for household 2

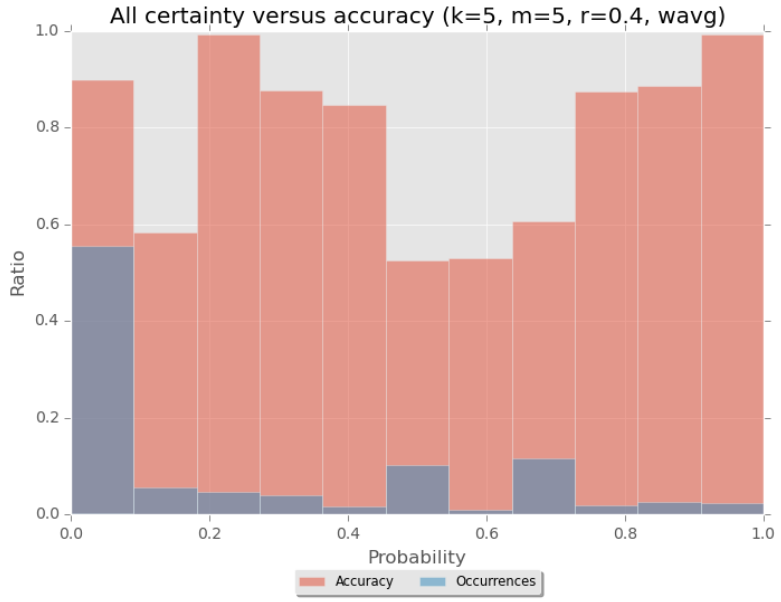

Figure 19: Fault tolerance certainty of predictions on household 2 ($r = 0.2, k = 5, m = 5, wavg$)

6.4 CONCLUSION

Varying results have been achieved in this experiment. For a lot of devices the algorithm is not capable of predicting that the device will be used in the near future. This is the matter for devices who do not have a strong relation with other devices. For a device to be predictable while his own sensor failed the device has to be on for more than 50% of the time when another device is on. A device which is barely used like a boiler or a dishwasher is simply not used often enough to fulfil this criteria. Other devices which are used often but not in a relation with other devices, such as a fridge and a freezer, are predicted to be used but is more a guess than a true prediction. The only really good predictions are made for the TV and stereo from household 2. The model shows that there is a really strong relation between those two devices, if the TV is on so is the stereo and vice versa. This tells us that the proposed approach works but it needs a strong relation between devices. For most devices this relation is not present and thus the proposed approach might not be the best solution for predicting device states in a household. The proposed method would work better with the combinations of devices and sensors. The results for the TV and stereo show us that the model probably would work for a combination of a light sensor, the time of the day

and a lamp. If the power measurements for the lamp fail the lamp his state can probably be predicted by looking at the light sensor, if it is dark the lamp is off and on if the room is illuminated during the night.

7 RELATED WORK

A lot of research has been done on smart buildings. This chapter provides an overview of the current state of the literature for each of the experiments performed in this thesis. Starting with related work for device (sensor) state prediction in smart buildings followed up by an overview of predicting energy consumption for buildings. After the energy prediction we will give insight in techniques available used in fault tolerant situations in sensor networks to recover erroneous measurements. Finally we take a short look at other applications for Markov models in the field of computing science.

7.1 STATE PREDICTION

There are multiple approaches to try to predict device usages. The most common approach is to find association rules between devices (sensors) which indicates what happens together and what not. Other publications use historical data of the appliance it self, and maybe other data, to predict whether it will be used again.

There are a couple of studies who really focus on whether a device will be used at a given time. This studies will be explained with some more detail than other publications. Other research will be discussed by categorizing the proposed approaches.

A PREDICTION SYSTEM FOR HOME APPLIANCE USAGE

A similar study has been done in 2013 by Basu et. al. in [12]. The proposed solution is to use learning algorithms with both expert knowledge and consumption data. The goal is to forecast whether an appliance while start in the next hour or not, but can also be used to predict for the next 24 hours. The forecasting data is restricted to the time of the event and the consumption data of the appliance to predict. As input for time there is the hour of the day (0-23), the season (1-4), the month (1-12) and the day of the week (1-7). Because the system uses the last 24 hours as input for forecasting the system is capable of learning from new data, it only needs 24 hours to learn a new device.

The study uses data from the Residential Monitoring to Decrease Energy Use and Carbon Emissions in Europe (REMODECE) project [6]. This project provides databases with characteristics of residential energy consumption in households. The data used by Basu et. al. is part of IRISE project which is a part of the REMODECE project and contains 10 minutes power consumption per appliance during one year in French households.

A decision tree, a decision table and a Bayesian Network (BN) are compared by looking at their accuracy for the predictions. Overall the results all have an accuracy over 90% but the results show that not each appliance is best predicted by one algorithm. Some devices work better with a decision table (lighting and an oven) but for the washing machine the decision tree gives the best results.

BARBATO 2011 The technique used in [11] works by predicting actual timeslots in which a device, probably, will be used. A big difference between Barbato and my approach is that my approach is making real-time predictions while Barbato predicts all the states the day before. Making the predictions on forehand can bring the advantage that you know how much electricity will be used the next day.

To make this predictions there are three steps:

1. *Status prediction.* During step 1 the algorithm looks at the last N days and based on that, and only that, will predict whether the device will be on at the next day. It works by denoting a string I where element j is either 1, indicating that the device was on on day j, or 0 meaning it hasn't been used. To predict the state for the device on the next day the probability that 1 or 0 is at the end of the new string. If the outcome is that the device will not be used the next day the algorithm stops here.

2. *Time prediction.* In step 2 the actual times where a device will be on is computed. For this purpose a normal function, with variance σ minutes, is used for each time the device was on in the last days. To forecast at what time the device will be on the sum of all normal functions is computed where a peak indicates that the device will be turned on on that time.

3. *Duration prediction* The last step in the process is to predict for how long a device will be used at time x after it has been turned on. To do this the duration of each time the device was on in step 2 is used. The duration on time x is the mean value of the durations whose normal functions contributed to the peak.

Results are reported for two use cases; a simulated data set and an experimental data set. The experimental data set contains an oven, a TV, a boiler and a computer and thus is similar to the devices used in this thesis. As Barbato et.al focus is not on exact prediction when a device will be used but more for how long during the day they don't report the same accuracy predictions as I did. What they do provide is an accuracy for step 1. They predicted correctly that the oven will be used with an accuracy of 76%, the TV with 82%, the boiler with 94% and the computer with 88%.

To compare the results of Barbato to my approach the algorithm has been implemented and is tested on the data set of household 2. The results per device for this experiment are listed in Table 12. The results show that that the proposed method is less capable of predicting when a device will be used than our proposed solution. This can be seen by comparing the results in Table 12 with the results in Table 4.

Statistic	Dishwasher	Air exhaust	Fridge	Freezer	Lamp	TV	Stereo
Correct	0,9854	0,962	0,5004	0,5146	0,8987	0,7156	0,6213
Correct off	0,9854	0,9923	0,6292	0,4954	0,9213	0,835	0,751
Correct on	0	0,0155	0,378	0,5281	0,4768	0,4159	0,448
Wrong	0,0147	0,0381	0,4997	0,4855	0,1014	0,2845	0,3788
Device off	0,9854	0,992	0,6256	0,4817	0,901	0,7635	0,6659
Device on	0,0147	0,0081	0,3745	0,5184	0,0991	0,2366	0,3342
On coverage	0	0,0598	0,5175	0,597	0,2446	0,501	0,5737
Off coverage	1	0,9693	0,4902	0,4259	0,9706	0,7821	0,6452

Table 12: Accuracy for Barbato et al on household 2

DISCOVERING USER BEHAVIOUR

A lot of studies focus on discovering user behaviours in smart environments. By discovering user behaviours it is possible to derive which devices they will use in a regular pattern. Aztiria et. al. [10] focus on discovering the user behaviour in the form of a sequence of actions. A follow up study in 2012 uses sensor data to discover frequent relationships between actions taken by the user and provides a language to represent them in [9]. A research done in 2011 by Rashidi et. al. introduced an automated approach for tracking activities in an individual his routine [42] without the need to have pre-selected activities and labeled training data.

ABNORMAL SEQUENCES

Some studies focus on abnormal sequences instead of reoccurring sequences.

The study in [37] tries to mine intertransaction association rules (IAR) to detect anomalous behaviour in smart environments. A similar goal is defined in [17] where the focus is on mining constraints for smart environments.

7.2 ENERGY CONSUMPTION

There are two methods for predicting energy consumption in a building: forward modelling and inverse modelling [22]. The most common approach is to use forward modelling which is based on an engineering model to predict energy consumption based on building geometry, weather data and materials used. Inverse modelling on the other hand uses a mathematical relationship between inputs and outputs. To limit the amount of related work we only look at inverse modelling research.

Next to the method to predict the energy consumption there is a difference based on the time interval that has to be predicted. There is short-term predicting which usually takes about one day to a week [34], from weeks till months is called medium-term predicting [25] and finally long-term predicting is related to yearly consumption [16].

Overall in the literature there are two dominant techniques on which the most research is based: Support Vector Machines (SVM) and Artifical Neural Networks (ANN), while the research mostly focuses on mid-term to long-term predictions for big office buildings.

SUPPORT VECTOR MACHINES

Recently a lot of research in forecasting building energy consumption is done using Support Vector Machines (SVM). The first study to use SVM for predicting energy consumption is done in [14]. The prediction of monthly energy consumption in a tropical region is performed on four buildings and reports results with a MSE in the range of 0,14-0,73. As an alternative Zhang et. al. proposed a method using a weighted SVM-model with a differential evolution algorithm in [50]. This method should be capable of predicting both short-term and mid-term energy. An experiment is made for a daily energy consumption prediction and a prediction for half-hourly energy consumption in a constitutional building in Singapore. Results are reported with a mean absolute percentage error (MAPE) of 5,843% for daily predictions and a MAPE of 3,767% for half-hourly predictions.

Jung et. al. proposed an approach by using the DSORCGA-LSSVM technique [30] for forecasting buildings daily energy usage. As LSSVM (Least Squared Support Vector Machine) requires the parameters to be selected accurate Jung uses DSORCGA, a hybrid of direct search optimization and real-coded genetic algorithm, to select those parameters. They report an average RMSE (root-mean-square error) varying from 7,5994 to 11,1319 depending on the data set used.

Parallel SVM is used in [40] to predict the energy consumption in multiple buildings. For this experiment all values are scaled to fit in the range $[0, 1]$ which results in a MSE (Mean Square Error) of $5.085e^{-05}$ and a SCC (squared correlation coefficient) of 0.9976.

ARTIFICAL NEURAL NETWORKS

Artificial Neural Networks (ANN) have been widely used for the use of non-linear time series to predict building energy consumption. Downside of an ANN is that the training is time consuming and that it can be trapped in a local minima [44]. ANN is used in [35] together with two adjustments to make hourly predictions. The first version takes an improved Particle Swarm Operation (iPSO) to adjust the ANN structure, weights and thresholds resulting in a MAPE of 0,0162. The second version uses a Genetic Algorithm-ANN and reports a MAPE of 0,0185. Finally a normal ANN achieves a MAPE of 0,0211.

Multiple methods have been compared in [45] where a decision tree, a neural network and regression tests are made and compared. Results are reported of a Root Average Squared Error (RSAE) of about 39kWh during the summer and 44-45kWh during winter for total week consumption.

HVAC SYSTEMS

A lot of research focuses on HVAC (heat, ventilation, air conditioning) systems energy consumption instead of predicting the loads from plugs. The prediction of heat consumption is done in [29] with the use of another ANN. In this study the results are compared for a feed forward backpropagation neural network (FFNN), a radial basis function network and a adaptive neuro-fuzzy interferences system (ANFIS).The results vary from a MAPE of 5,25% to 5,79% for FFNN, 5,43% to 6,01% for RBFN and 5,43% to 5,64% for ANFIS. Cooling consumption is predicted in [36] with the use of a back propagation neural network (BPNN), a radial basis function neural network (RBFNN), a general regression neural network (GRNN) and finally a support vector machine (SVM).

An annual occupant based energy consumption prediction was made in [47] using a combination of a Markov chain and the Monte Carlo method. The error rate varies from 0,99% to 3,95% depending on the office to predict.

COMPARISON The techniques found in the literature are compared in Table 13 which shows for each publication the used technique and the reported results. As not each study uses the same error measurement and same data it is hard to compare results directly to each other but it does give an indication.

Source	Technique	Results
[50]	weighted-SVM	MPAE 3,767
[30]	DSORCGA-LSSVM	RMSE 7,5994-11,1319
[40]	Parallel SVM	SCC 0.9976
[35]	iPSO-ANN	MAPE 0,0162
[35]	GA-ANN	MAPE 0,0185
[35]	ANN	MAPE 0,0211
[29]	FFNN	MAPE 5,25%-5,79%
[29]	RBFN	MAPE 5,43%-6,01%
[29]	ANFIS	MAPE 5,43%-5,64%
[36]	BPNN	RMSE 2,365%
[36]	RBFNN	RMSE 1,507%
[36]	GRNN	RMSE 1,201%
[36]	SVM	RMSE 1,182%
[47]	Markov-Monte Carlo	0,99%-3,95%
[14]	SVM	MSE 0,14-0,73
[45]	DT, ANN, regression	RASE 39kWh and 45kWh

Table 13: Comparison of different energy prediction methods from the literature

7.3 FAULT TOLERANCE

A lot of research is conducted in the field of fault tolerance for sensors and especially for wireless sensor networks (WSN). From this research a significant part focuses on deploying sensors in a fault tolerant matter. Sensor validation could be done by either hardware redundancy, analytical redundancy or temporal redundancy [24]. For this literature study we ignore hardware redundancy and look at research that identifies sensor failure and reconstructing the erroneous measurements.

Research to identify faulty sensor measurements in [8] describes an approach to estimate the value of one variable based on the value of correlated variables. This is similar to our approach where we try to predict the state of a device based on the state of other devices.

Bayesian networks have been used for representing relationships between variables in [27]. The proposed model utilizes probabilistic propagation to the expected value of variables. When the expected value differs from the observed value a fault is assumed. A method using a Bayesian belief network (BNN) is proposed in 2004 by Merhanbod et. al [38].

In 1991 a method is described to detect sensor failure and recover the correct value by using an ANN [26]. The goal of this research is to train a neural network which should identify that one sensor his output does not match with other sensors their output. A second neural network is trained to retrieve the correct value of the faulty sensor. A method for detection the failure of multiple sensors is presented in [43]. The discussed method uses a neural network based sensor fault-detection, isolation and accommodation (SFDIA) scheme to detect consecutively sensor failure by taking an analytical redundancy approach.

Principal Component Analysis (PCA) [28] is widely used in the literature for fault detection and reconstruction [49, 20, 32].Principal Component Analysis can be used to build a model that characterises the correlation between sensors in a system. Such an model is used to reconstruct faulty sensor measurements in [46].

7.4 OTHER USAGES OF MARKOV MODELS

Markov models are widely used for all kind off applications. In the field of computing science Markov models have been used by Whittaker et. al. for statistical software testing [48], used to create web page recommendations together with page ranking by Eirinaki et. al in [23], applied for image classification by Dilligenti et. al in [19] and was already used in 1998 to perform face detection and recognition in [39].

In the context of smart buildings and environments Markov models have been used to mine a user context based on location time series in 2012 [31]. Duong et. al used Markov models for learning and recognizing human activities of daily living (ADL) which is important in a smart environment [21].

8 CONCLUSION AND FUTURE WORK

In this thesis we tried to find an answer for the research question: *'How can patterns be mined for household appliances by monitoring electric plug loads?'*. To answer this question three sub-questions were formulated and for each question an experiment was conducted. Now we will try to answer each of the sub-questions and use those answers to formulate a conclusion for the research question. After formulating a conclusion insight will be given in possible future work to improve the proposed solution.

CAN A DEVICE STATE BE FORECASTED BY MONITORING PLUG LOADS?

In Chapter 4 an experiment is conducted to find out if device states can be predicted by monitoring plug loads in a building. The results show that on average the state of devices is predictable but there is a difference between devices that are used often and devices that are barely used. Devices that are barely used are hard, if not impossible, to predict that they will be used while devices that are used often, in a regular pattern, reach high accuracy.

CAN A DEVICE HIS UPCOMING ENERGY CONSUMPTION BE PREDICTED BY MONITORING PLUG LOADS?

The experiment performed in Chapter 5 reports similar results as the previous question. As this experiment used results from the previous experiment errors from the state prediction are reflected in this experiment. The results show that appliances who have a steady consumption profile can be predicted accurate. Devices like a TV, lights, dryer and a dishwasher have a regular consumption profile and this can be predicted with high accuracy. Devices which work more with extreme peaks, e.g. a boiler, a printer or a fridge, are harder to predict. Overall the energy prediction is following the actual consumption rather than forecasting it.

CAN A DEVICE HIS STATE STILL BE FORECASTED BY MONITORING PLUG LOADS ONCE YOU STOP MONITORING THE DEVICE ITSELF?

The experiment, done in Chapter 6, shows that it is still possible to predict a device his state once you stopped monitoring the device. This requires the device to have a strong relationship with other device(s). The experiment showed that the TV and the stereo are on together almost during 100% of the time. This tells that after disabling the sensor for either of those it is still possible to predict the other device his state based on the remaining device. Devices that are barely used are impossible to predict in this case as their prediction relies highly on themselves.

HOW CAN PATTERNS BE MINED FOR HOUSEHOLD APPLIANCES BY MONITORING ELECTRIC PLUG LOADS?

The results of the experiments show that Markov models are capable of finding sequences in the usage of household appliances. The experiments have shown that for a lot of appliances it is possible to accurately predict when they will be used and how much energy it will consume. The accuracy of the predictions depends highly on the amount of time the appliance is used and if it happens in regular patterns or on random moments. The model works well on devices which are used often and achieves accuracy of close to 100% in state prediction for some devices. At the same time those devices have a mean absolute error of only a few Watt when predicting their energy consumption. This makes the model reliable on those devices. In the meanwhile there are also devices who are really hard to predict and achieve an accuracy of near 100% but only achieve that accuracy because the

device is barely ever used so it is predicted off all the time. Besides rarely used devices the devices that run in a cyclic pattern are found to be hard to predict too. Fridges and freezers are activated every x minutes to keep the temperature at a certain threshold and are not triggered by user actions. As those devices are not used in a sequence with other devices the proposed model is not the best solution for those kind of devices. The patterns mined show interesting results and probably can be used in further research to actually reduce the energy consumption in households or other applications.

8.1 FUTURE WORK

There are a couple of interesting points to continue this work on. First of all this thesis mines patterns for household appliances in households. The actual usability of those patterns has to be validated in a real household to verify if energy can be saved. A household that is not connected to the grid would be a good experiment to test this model on. A household with limited energy available, for example only by solar panels, can benefit from forecasting energy consumption. As an example an energy management system can decide to not start the washing machine if the model predicts that the consumed energy will be higher than the generated energy if the washing machine also will be used.

Secondly the complexity of the proposed model is high. With a dense data set of 10 devices the complexity of the model increases so fast that it is already very time consuming to train the model. In this thesis an optimization technique is applied by pruning infrequent sequences from the model which reduces the complexity of the model significantly. The downfall of pruning the model is that devices which are used less often will be removed from the model and are thereby harder or even impossible to predict. Another limitation of the pruning technique is that it makes it harder to periodically increment the model. Because some data is removed from the model it is impossible to add new observations to it. For this reason it is also hard to add new devices to the model. New devices will be observed as a barely used device and will be pruned from the model.

Furthermore it might be interesting to combine plug loads with other sensors in the model. Including light sensors, occupancy sensors or pressure sensors might increase the accuracy of the model.

A

DETAILED RESULTS FOR STATE PREDICTION

In this appendix the state prediction experiment is validated in detail by looking at device specific results for each data set. This is done by first looking at the most supported sequences per device followed by an interpretation of the achieved results.

A.1 RUG DATA SET

To validate the RUG data set we look first at the supported sequences per device followed by the device specific results.

A.1.1 *Supported sequences per device*

In this section the top 5 most supported sequences per device are reported. Those sequences should give an insight in correlation between devices, if any. All devices are reported in a table with in the first column the sequence, in the second column how often it occurred in the data set and in the last column the probability that the given device will be on in the next m time-units. The data is taken from the trained database with the configuration $k = 5$, $m = 5$, $r = 0.2$.

TV-SCREEN The most supported sequences for the TV-screen (Table 14) indicate that the screen is an independent device. The support for the screen is really high (near 100%) when the screen is already in the sequence. There seems to be almost no relation between the TV-screen and other devices. The only sequence in the five most supported sequences that contains another appliance than the screen is the least supported sequence. This sequence, containing the boiler and the printer, however only has a probability of 17.3% that the screen will be on if the sequence is present.

Sequence	**Occurrences**	**Support**
Screen ⇒ Screen ⇒ Screen ⇒ Screen	17976	0.998
Screen ⇒ Screen ⇒ Screen	17992	0.997
Screen ⇒ Screen	18008	0.996
Screen	18024	0.996
Boiler ⇒ Printer ⇒ Printer ⇒ Boiler	36	0.173

Table 14: Top 5 supported sequences for the TV-screen in the RUG data set

MICROWAVE The microwave shows a strong relationship between the microwave and the coffee maker and printer as can been seen in Table 15. In all five sequences the coffee maker and/or the printer is present. This suggests that people tend to put the microwave on and while their food is being prepared they decide to get some coffee or print something. The fact that all sequences have a support over 80% indicates that the microwave is an appliance which should be able to predicted reliable.

BOILER Following the data in Table 16 there seems to be a relation between the boiler and the coffee machine and the printer. Just as with the microwave are the coffee maker and/or printer present in all the sequences for the boiler. This might suggest that people tend to get coffee and print something while getting hot water. Another option is that the coffee machine uses hot water from the boiler but this sounds unlikely. An important note is that even

Sequence	Occurrences	Support
Microwave ⇒ Coffee maker ⇒ Printer ⇒ Microwave	110	0.875
Microwave ⇒ Printer ⇒ Printer ⇒ Microwave	770	0.828
Microwave ⇒ Printer ⇒ Microwave ⇒ Microwave	879	0.822
Microwave ⇒ Coffee maker ⇒ Coffee maker ⇒ Microwave	154	0.822
Microwave ⇒ Microwave ⇒ Printer ⇒ Microwave	872	0.819

Table 15: Top 5 supported sequences for the microwave in the RUG data set

the best supported sequence containing the boiler has only a support of 32.6% which shows that the boiler is not used often. As the best supported sequence only has a support of 32% it is impossible for the algorithm to ever predict the boiler in the on-state. As a result the prediction for the boiler will always be off and the accuracy equals the times the boiler is off.

Sequence	Occurrences	Support
Boiler ⇒ Coffee maker ⇒ Coffee maker ⇒ Boiler	33	0.326
Boiler ⇒ Coffee maker ⇒ Boiler ⇒ Boiler	66	0.313
Boiler ⇒ Printer ⇒ Printer ⇒ Boiler	118	0.309
Boiler ⇒ Boiler ⇒ Printer ⇒ Boiler	137	0.299
Boiler ⇒ Coffee maker ⇒ Coffee maker ⇒ Coffee maker	13	0.279

Table 16: Top 5 supported sequences for the boiler in the RUG data set

PRINTER Table 17 represents the five most supported sequences for the printer. The supported sequences show a strong relationship between the printer and the coffee maker. With 3 out of 5 of the sequences containing the coffee maker and the support being higher than 85%. This indicates, as stated before with the microwave, that people tend to get coffee or prepare their food if they are waiting for the printer to finish. This theory is backed by the last sequence where people start with putting their food in the microwave and print something afterwards.

Sequence	Occurrences	Support
Printer ⇒ Coffee maker ⇒ Coffee maker ⇒ Printer	268	0.929
Printer ⇒ Coffe maker ⇒ Printer ⇒ Printer	663	0.926
Printer ⇒ Printer ⇒ Printer ⇒ Printer	7386	0.884
Printer ⇒ Coffee maker ⇒ Printer	1140	0.847
Microwave ⇒ Printer ⇒ Microwave ⇒ Printer	760	0.842

Table 17: Top 5 supported sequences for the printer in the RUG data set

COFFEE MAKER The data in Table 18 shows an interesting relation between the boiler and the coffee maker. According to those sequences the boiler seems to be active before people use the coffee machine. Another interesting observation is that the coffee maker tends to be running for a longer period at a time, not just for one cup of coffee. Showing that the top four sequences don't actually start with the coffee machine can indicate that taking coffee is a single task, unlike for example heating food in the microwave. Just as with the boiler the coffee maker also doesn't have a single sequence whose support is over 0.5 meaning that it is impossible to predict the on state accurate.

A.1.2 *Device specific validations*

Now we have the results for all devices together it is interesting to look at the values for individual devices. Is one device his behaviour more predictable

Sequence	Occurrences	Support
Boiler ⇒ Coffee maker ⇒ Coffee maker ⇒ Coffee maker	15	0.459
Microwave ⇒ Microwave ⇒ Coffee maker	442	0.347
Boiler ⇒ Printer ⇒ Boiler ⇒ Printer	16	0.332
Boiler ⇒ Boiler ⇒ Printer ⇒ Printer	28	0.328
Coffee maker ⇒ Coffee maker ⇒ Coffee maker ⇒ Coffee maker	28	0.327

Table 18: Top 5 supported sequences for the coffee maker in the RUG data set

than another device? As stated before we have five devices: a microwave, a TV-screen, a printer, a boiler and a coffee machine. For each device we will do the same as with the optimal results for all devices together, look at the success rate and go deeper into the numbers by plotting a histogram with the certainty and accuracy. The devices individual results are listed in Table 19 and will be used in the discussion of the individual devices.

Statistic	Screen	Microwave	Boiler	Coffee maker	Printer
Correct	0,9998	0,967	0,9691	0,8077	0,7453
Correct off	0,9998	0,9734	0,9691	0,8077	0,79
Correct on	0,9334	0,9084	0	0	0,4588
Wrong	0,0003	0,0331	0,031	0,1924	0,2548
Device off	0,9996	0,8861	0,9691	0,8077	0,7564
Device on	0,0005	0,114	0,031	0,1924	0,2437
On coverage	0,3889	0,7895	0	0	0,2544
Off coverage	1	0,9898	1	1	0,9034

Table 19: State prediction results for the RUG data set

MICROWAVE We will start by looking at the values for the microwave in Table 19. As can be seen in the table the results for the microwave on his own are better than the average. The predictions of the microwave are $\approx$ 97% accurate while the predictions that the microwave is on are $\approx$ 90% correct and off is more than 97% accurate. The coverage shows even better results with an on coverage of almost 80% and an off coverage of 99%.

The histogram in Figure 20 clarifies those results. By looking at the histogram you can see that there are barely any uncertain predictions with a certainty of near 0.5. More than 80% of all predictions have a certainty lower than 0.2 which gives a near 100% accuracy. An interesting view is provided by looking at the certainties greater than 0.5, certainties where the microwave will be predicted as on. There are no predictions with a certainty over 0.7 yet the accuracy is still over 90%.

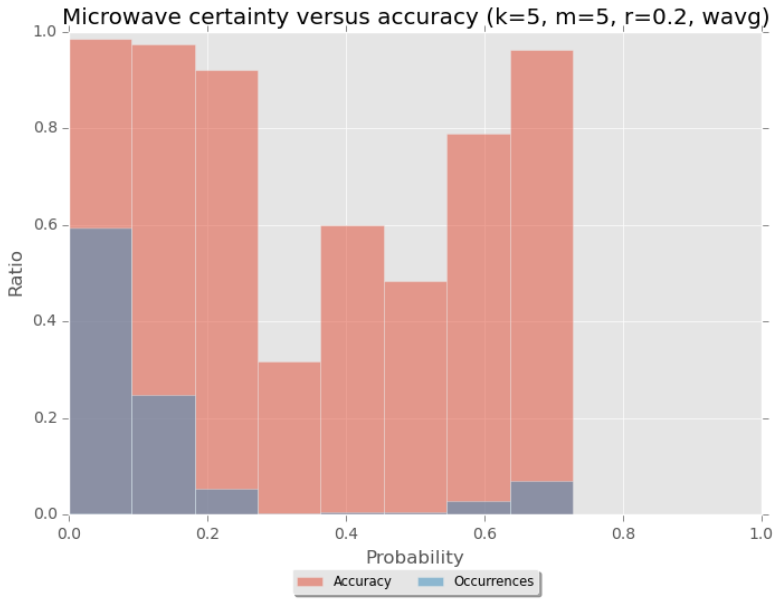

Figure 20: Histogram of certainty of predictions for the microwave in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

PRINTER The printer is another device which is really user triggered. The results in Table 19 indicate that the usage of the printer is a bit harder to predict with an accuracy of $\approx 75\%$. The difference between the off coverage and the correct off accuracy indicates that the printer is too often to be predicted off while it actually is going to be on.

By looking at the data in the histogram (Figure 21) the cause of those results are clearly visible. There are a lot of predictions where the certainty is 0.0, around 60% of the predictions. The most of the remaining predictions have a certainty between 0.4 and 0.6. This is the area where the algorithm is not certain whether a device will be used or not causing, relatively, a lot of errors. A weird observation is that there are no predictions with a certainty of 0.1, 0.2 or higher than 0.7. So the algorithm is either really sure that the printer will not be used with a certainty of 0.0, but still only gaining an 80% accuracy, or has an rough idea with almost the same accuracy.

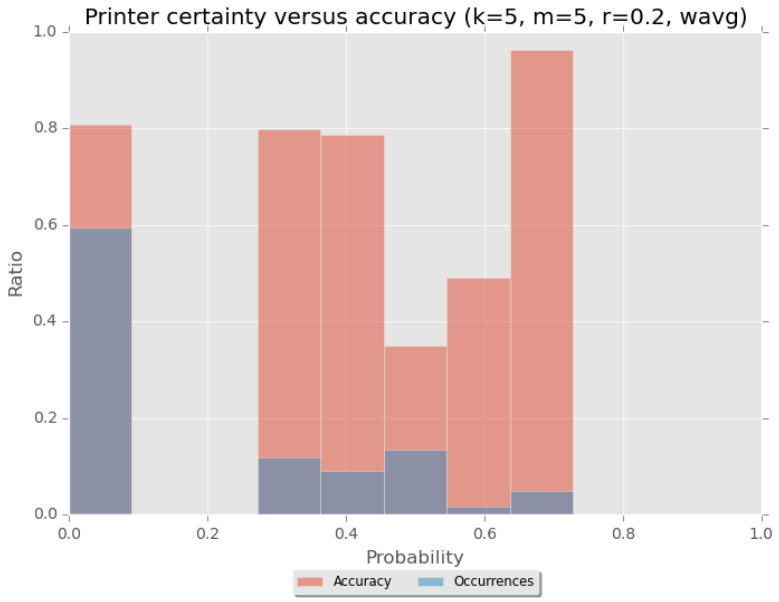

Figure 21: Histogram of certainty of predictions for the printer in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

BOILER The boiler should be a device which operates under predictable patterns as the water has to be reheated every once in a while.

The results in Table 19 at first sight show that this assumption is correct: the accuracy of the predictions is almost 97%. However when looking at the accuracy of predicting that the boiler will be on the accuracy is zero. This is caused by the fact that the boiler is only on during 3% of the time, or a total of 2600 minutes of the over 86000 minutes. This results in a histogram where the algorithm is 80% of the time certain that the boiler will not be used as can be seen in Figure 22. Mathematical this is an expected result. Looking at the restriction rate of 0.2, meaning that only sequences which occur 20% of the times are supported, it is impossible for the boiler to occur in sequences with a high probability. Therefore there are no situations where the algorithm would predict the boiler as on and thus the accuracy for device on is zero. Concluding that a boiler is a predictable device but the only reason for this is that it is off almost all the time. Looking at the results for the boiler with $r = 0$ gives the same results, the boiler is just not on often enough to reliable predict when it will be on. The accuracy with a certainty of 0.2 implicates that the boiler is actually quite often on at this stage. A possible solution for this might to introduce device dependent certainty threshold (e.g. suggest the boiler to be used with a certainty of 0.25).

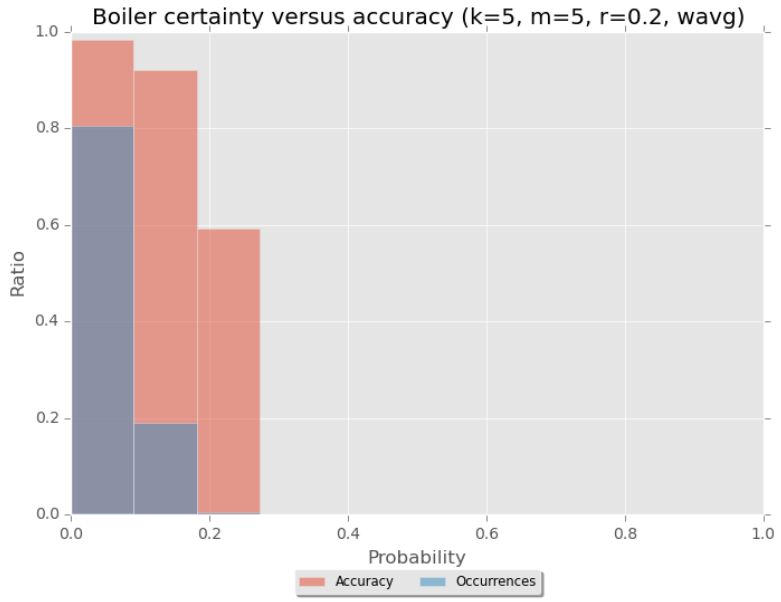

Figure 22: Histogram of certainty of predictions for the boiler in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

TV-SCREEN By looking at the raw transaction data for the validation set it shows that TV-Screen has only be on for 36 minutes in total. Looking at the training set shows that the TV-screen there was only on for about two weeks in a row. This results in the change of the TV being on after it was on to a stunning 99%. As a result this is not a really representative result for the TV-Screen. The statistics as in Table 19 show an accuracy of 99,9998% which is due to the fact that the TV-screen was 99,9996% of the time off. Interesting is that the accuracy of predicting the TV-screen on is still a good 93%. This can be easily explained with the data from Figure 23. It shows that the accuracy of predicting the device on is 100% while the certainty is bigger than 0.6. The reason why the algorithm is so certain that the TV-screen is on is also easy explained. As soon as the TV-Screen is on the training set indicates that the TV-screen will stay on (99% of the time). As a result the TV-Screen will be predicted to stay on when the screen is on and to stay off once the screen is off. The actual transition from on to off or the other way around will not be noticed and thus not predicted correctly. Examining the off coverage from Table 19 shows that the algorithm only once didn't predict the screen to be off while it actually was off (86024 times versus 86025 times). Although this results are not superb, due to the limited on time, it shows that the screen is actually really great to predict.

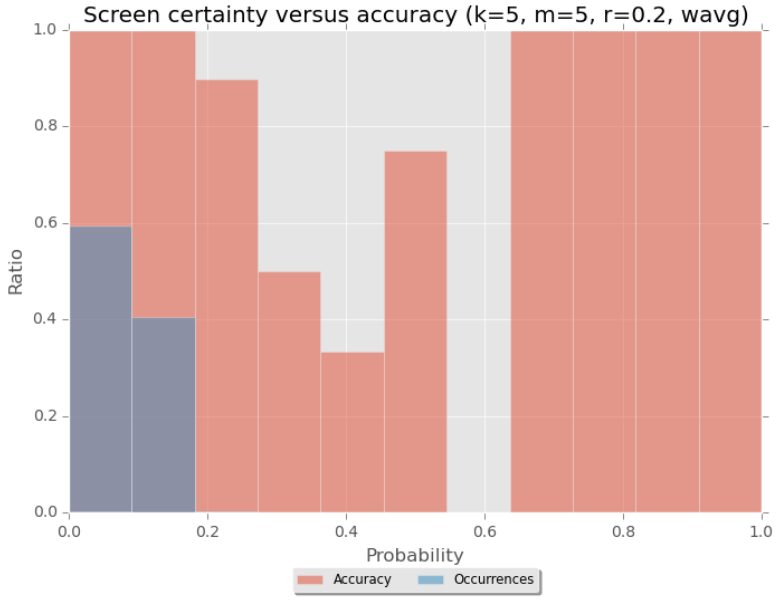

Figure 23: Histogram of certainty of predictions for the screen in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

COFFEE MACHINE The last device to be investigated is the coffee machine. At first glance, without looking at actual data, the coffee machine seems to be an unpredictable machine as people tend to walk by at random moments and take a coffee, apart from breaks where a peek in consumption makes sense. The results in Table 19 seems to support my theorem. Although the coffee maker achieves an accuracy of 80% the situation is the same as the previously mentioned boiler. There is simply no sequence whose support is bigger than 0.5 and as a result the coffee machine is never predicted on. A quick glance at the histogram in Figure 24 shows that the algorithm is uncertain about the state of the coffee machine almost all the time. The certainty that the coffee machine will be used is 0.4 for over 80% of the time.

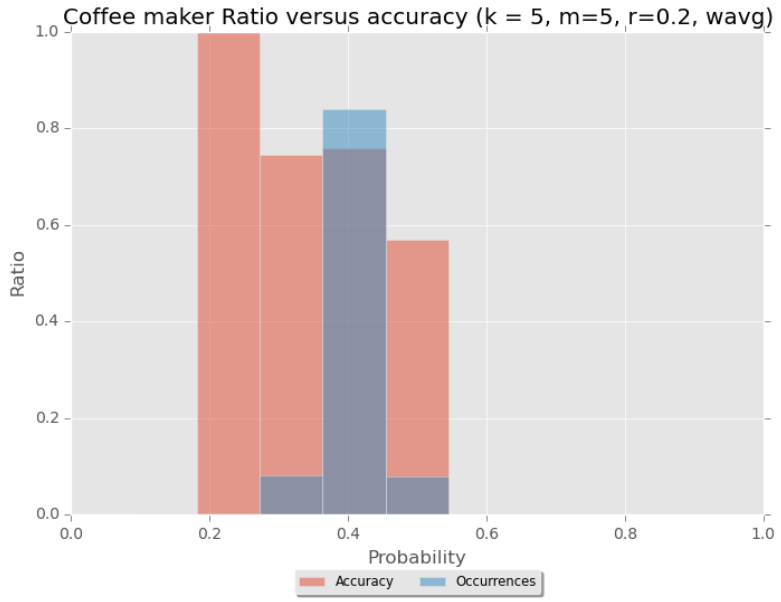

Figure 24: Histogram of certainty of predictions for the coffee machine in the RUG data set ($r = 0.2, k = 5, m = 5, wavg$)

A.2 DETAILED RESULTS FOR HOUSEHOLD 1

The results for household 1 will be explored in the same was as for the Living Lab Study data set by first looking at the supported sequences and then the device specific results.

A.2.1 *Supported sequences per device*

The data is taken from the trained database with the configuration $k = 5$, $m = 5$, $r = 0.2$.

FRIDGE The best supported sequences for the fridge are listed in Table 20. This list shows that for the fridge there are some sequences with a really high probability that the fridge will be used in the near future. An interesting observation is that sequences which include the freezer occur way more often than sequences with the washing machine or the dryer (45000 occurrences versus a few thousand occurrences). This indicates that the fridge is more often on around the same time the freezer is on than with the dryer or washing machine. The most supported sequence that doesn't contain the fridge it self has a support of 0.511, to compare the least supported sequence which still contains the fridge has a support of 0.856. This indicates that the fridge is only, reliable, predicted on when the fridge was on recently.

Sequence	Occurrences	Support
Fridge ⇒ Feezer ⇒ Fridge ⇒ Fridge	45990	0.964
Washing machine ⇒ Fridge ⇒ Fridge ⇒ Fridge	5996	0.963
Fridge ⇒ Fridge ⇒ Freezer ⇒ Fridge	45973	0.963
Dryer ⇒ Fridge ⇒ Freezer ⇒ Fridge	1805	0.963
Freezer ⇒ Fridge ⇒ Fridge ⇒ Fridge	46915	0.963

Table 20: Top 5 supported sequences for the fridge in household 1

FREEZER The freezer doesn't seems to have a strong 'on if on' relationship as can be seen in Table 21. For the top five supported sequences all sequences end with the freezer but none of the five sequences contains the freezer in another part of the chain. This observation is partially confirmed by the sequences only containing the freezer. With the sequence {Freezer} the change of the freezer being on in the near future is 'only' 80%. This model suggest that the freezer is better predictable while other devices are present.

Sequence	Occurrences	Support
Washing machine ⇒ Wasching machine ⇒ Fridge ⇒ Freezer	3557	0.926
Washing machine ⇒ Fridge ⇒ Washing machine ⇒ Freezer	3539	0.926
Washing machine ⇒ Fridge ⇒ Fridge ⇒ Freezer	3601	0.926
Dryer ⇒ Dryer ⇒ Fridge ⇒ Freezer	1853	0.926
Dryer ⇒ Fridge ⇒ Fridge ⇒ Freezer	1815	0.925

Table 21: Top 5 supported sequences for the freezer in household 1

WASHING MACHINE For the washing machine the model suggest a strong independent relation for the fridge. As can be seen in Table 22 the washing machine tends to be on when the washing machine already was on. At the same time the fridge shows up often in the sequences which might indicate a relationship between the washing machine and the fridge. This is not a pattern you would manually define so it can be interesting to find out why this happens. A possible, and likely, explanation is that the fridge is the most used device in this data set which causes it to show up in sequences more often.

Sequence	Occurrences	Support
washing ⇒ washing ⇒ Fridge ⇒ washing	6027	0.982
washing ⇒ Fridge ⇒ washing ⇒ washing	6355	0.981
washing ⇒ Fridge ⇒ Fridge ⇒ washing	5797	0.981
washing ⇒ Freezer ⇒ Fridge ⇒ washing	3569	0.981
washing ⇒ Fridge ⇒ Freezer ⇒ washing	3603	0.981

Table 22: Top 5 supported sequences for the washing machine in household 1

DRYER For the dryer you manually would expect a relationship between the washing machine and the dryer. This is because usually as soon as the washing machine finishes you put your wet laundry in the dryer to dry. The aforementioned relationship isn't supported by the model as can be seen in Table 23. Just as with the washing machine the dryer is mainly supported by sequences where the dryer is already present.

Sequence	Occurrences	Support
Dryer ⇒ Fridge ⇒ Freezer ⇒ Dryer	1902	0.990
Dryer ⇒ Fridge ⇒ Dryer ⇒ Dryer	3148	0.989
Dryer ⇒ Fridge ⇒ Fridge ⇒ Dryer	3052	0.989
Dryer ⇒ Freezer ⇒ Fridge ⇒ Dryer	1893	0.989
Dryer ⇒ Dryer ⇒ Fridge ⇒ Dryer	3137	0.989

Table 23: Top 5 supported sequences for the dryer in household 1

A.2.2 *Device specific validations*

In Table 24 the same data as in Table 3 is listed for convenience. In this section for each of the devices an attempt is made to explain the results by looking at the certainty of the predictions per device.

Statistic	Fridge	Dryer	Washing machine	Freezer
Correct	0,8764	0,9922	0,971	0,7173
Correct off	0,9082	0,9928	0,9749	0,7198
Correct on	0,8303	0,9738	0,9209	0,7167
Wrong	0,1237	0,0079	0,0291	0,2828
Device off	0,6063	0,9606	0,9104	0,3806
Device on	0,3938	0,0395	0,0897	0,6195
On coverage	0,8621	0,8238	0,7397	0,8993
Off coverage	0,8856	0,9991	0,9938	0,4213

Table 24: State prediction results for household 1

FRIDGE The certainties of the predictions for the fridge reveal that it is difficult to predict the state of the fridge. As can be seen in Figure 25 near 40% of all the predictions has a certainty of 0.5 which indicates it is an estimated guess which, surprisingly, still achieves an accuracy over 90%. Another observation is that the algorithm is either totally sure that the fridge will stay off, certainty 0.0, or is 0.5 or above. This abnormality can be explained by the supported sequences in Table 20 which tells that the fridge will not be used if it wasn't on recently.

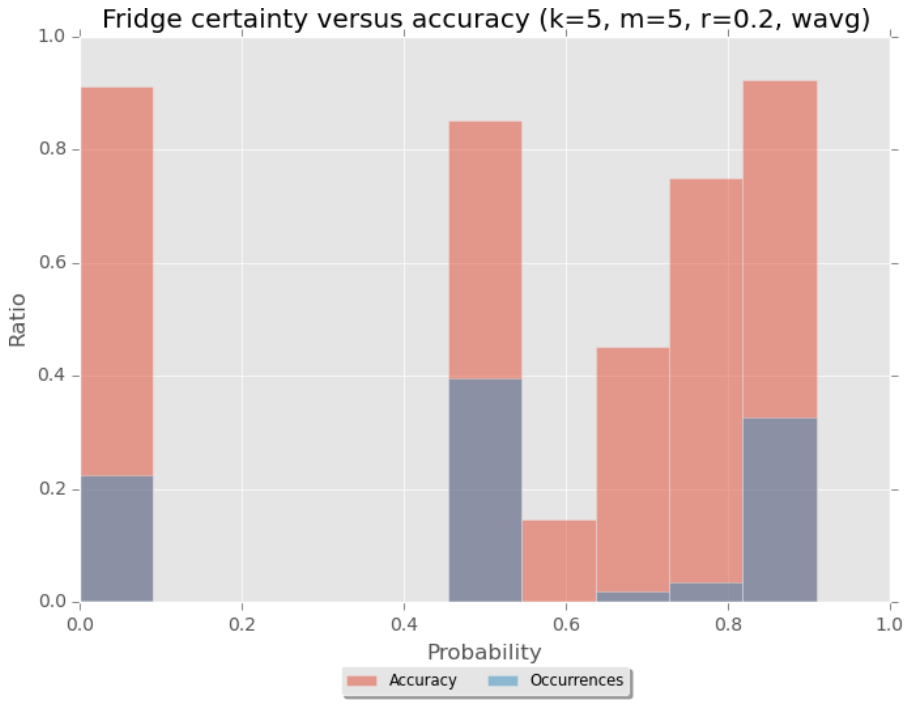

Figure 25: Certainty of predictions for the fridge in ECO household 1 ($r = 0.2, k = 5, m = 5, wavg$)

FREEZER The distribution of the certainties for the freezer suggest a reliable prediction, all predictions are either with a certainty of 0.0 or 0.7 or higher. The accuracy tells a different story with an accuracy of not even 30% with a certainty of 0.7. It is hard to explain this drop in accuracy as the model shows that each supported sequence has a support for the freezer of at least 67%, also when the freezer it self is not in the sequence. Meanwhile this observation explains the weird behaviour inFigure 12.

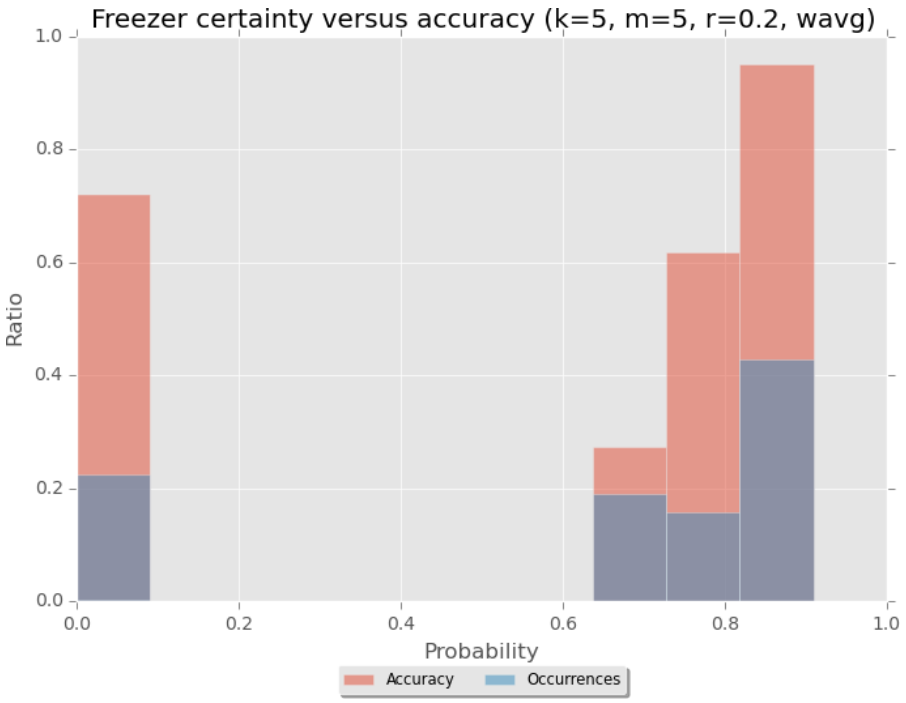

Figure 26: Histogram of certainty of predictions for the freezer ($r = 0.2, k = 5, m = 5, wavg$)

WASHING MACHINE The washing machine shows some interesting results. About 90% of all the predictions are made with a certainty of 0.0 or 0.1 with both accuracies close to 100%. Although there are barely any predictions with other certainties the predictions made with those certainties all gain a suitable accuracy too, ranging from $\approx$ 70% to almost 100%. Only the predictions with a certainty of 0.9, as shown in Figure 27, show a low accuracy of around 60%. Just as with the fridge this behaviour cannot be explained by looking at the model and it remains unclear why this achieves only an accuracy of 60%.

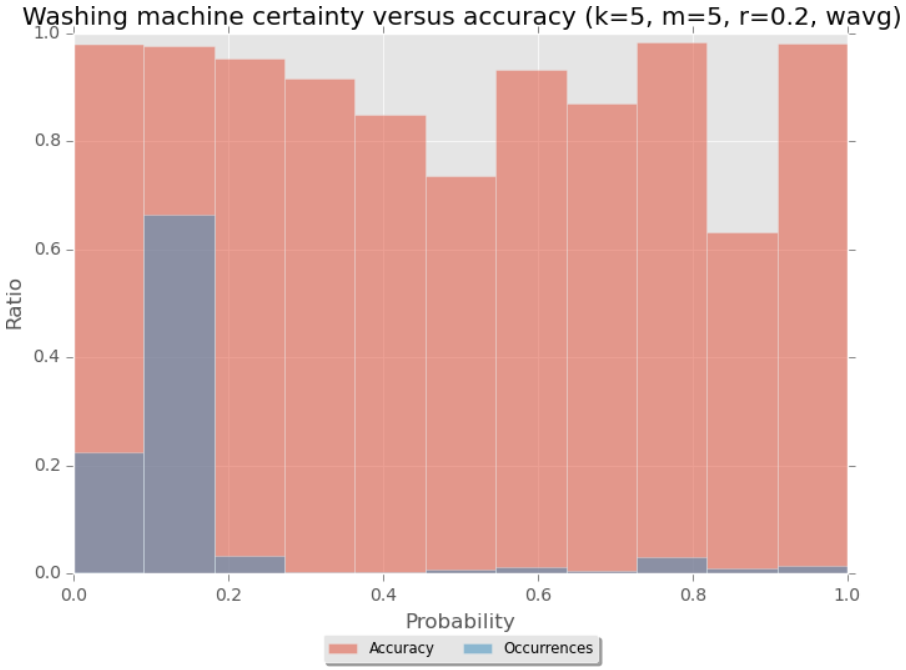

Figure 27: Histogram of certainty of predictions for the washing machine ($r = 0.2, k = 5, m = 5, wavg$)

DRYER The results for the dryer are as you expect for most of the devices. The graph in Figure 28 shows a parabola like function where the accuracies are lower near 0.5 and higher the further away from 0.5.

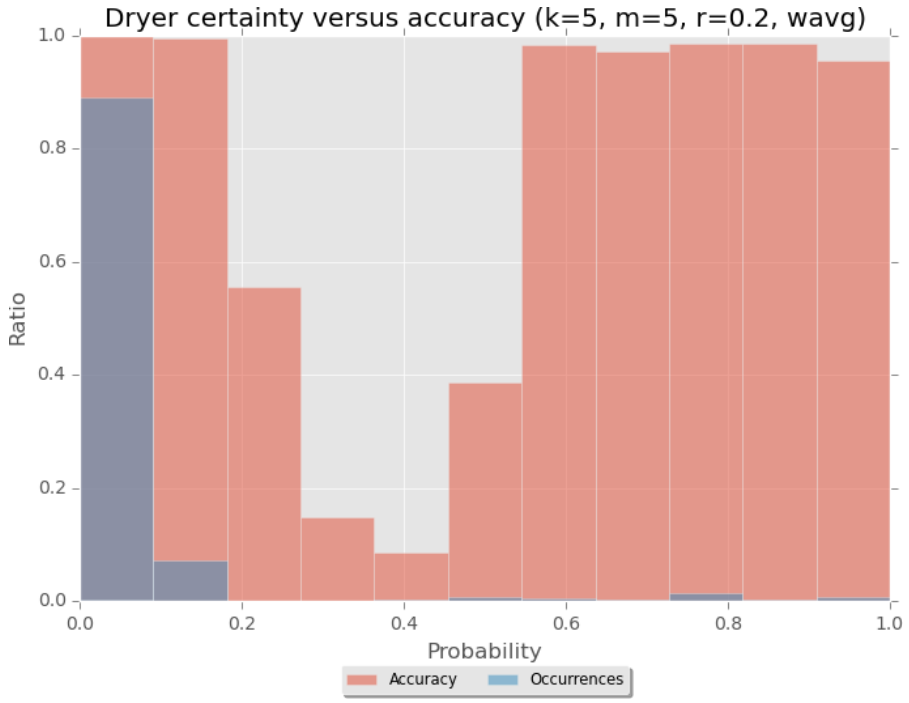

Figure 28: Histogram of certainty of predictions for the dryer ($r = 0.2, k = 5, m = 5, wavg$)

A.3 DETAILED RESULTS FOR HOUSEHOLD 2

The approach for household 2 is the same as with the previous data sets. The difference with this data set and the other data sets is that this data set is configured with a support threshold of 0.4. The reason for this is that with $r = 0.2$, as the other data sets, there would be too much supported sequences.

A.3.1 *Supported sequences per device*

The data is taken from the trained database with the configuration $k = 5$, $m = 5$, $r = 0.4$.

DISHWASHER A dishwasher seems to be a device which is mostly related to itself. For all the sequences which contain the dishwasher, so not just the top 5 from Table 25, the support is at least 90%. As soon as the dishwasher is not in the sequence the support drops to 7% at best. This is not an unexpected result as the dishwasher is only used about 1.5% of the total time. This means that other frequent sequences have a really small change of having impact on the dishwasher.

Sequence	Occurrences	Support
Dishwasher ⇒ Stereo ⇒ Dishwaser ⇒ Dishwaser	1638	0.978
Dishwasher ⇒ Dishwaser ⇒ Stereo ⇒ Dishwasher	1627	0.978
Dishwasher ⇒ Stereo ⇒ Stereo ⇒ Dishwasher	1631	0.978
Dishwasher ⇒ TV ⇒ Dishwasher ⇒ Dishwasher	1199	0.976
Dishwasher ⇒ Dishwaser ⇒ TV ⇒ Dishwasher	1195	0.976

Table 25: Top 5 supported sequences for the dishwasher in household 2

AIR EXHAUST For the air exhaust the same rules apply as for the dishwasher as the air exhaust is only used for a few minutes every day, probably while cooking. A difference with the dishwasher that the support for the air exhaust is at most around 80%, as Table 26 shows, while the air exhaust is used. Although this shows a strong relation it is by far not as trustworthy as other devices which reported a support near 100% when the device itself was in the sequence. The most supported sequences which doesn't involve the air exhaust has only a support of 3% which tells us that the air exhaust simply is off while it wasn't on recently.

Sequence	Occurrences	Support
Air exhaust ⇒ Stereo ⇒ Air exhaust ⇒ Air exhaust	559	0.827
Air exhaust ⇒ Stereo ⇒ Stereo ⇒ Air exhaust	589	0.827
Air exhaust ⇒ Stereo ⇒ Fridge ⇒ Air exhaust	359	0.825
Air exhaust ⇒ Air exhaust ⇒ Stereo ⇒ Air exhaust	591	0.825
Air exhaust ⇒ Air exhaust ⇒ TV ⇒ Air exhaust	326	0.821

Table 26: Top 5 supported sequences for the air exhaust in household 2

FRIDGE The data in Table 27 shows an interesting relation between the fridge and the air exhaust. Although the air exhaust is barely used it shows that when the air exhaust is on the fridge tends to be used also. A logical explanation is that people put the air exhaust on as they start cooking and get their food out of the fridge when the pan is hot. In the meanwhile there are a total 648 supported sequences whose probability is over 50% that the fridge will be used if that sequence occurs with a total number of supported sequences of 652.

Sequence	Occurrences	Support
Air exhaust ⇒ Stereo ⇒ Fridge ⇒ Fridge	434	0.909
Air exhaust ⇒ Fridge ⇒ Air exhaust ⇒ Fridge	422	0.907
Air exhaust ⇒ Fridge ⇒ TV ⇒ Fridge	249	0.906
Air exhaust ⇒ Air exhaust ⇒ Fridge ⇒ Fridge	541	0.904
Air exhaust ⇒ Fridge ⇒ Fridge ⇒ Fridge	552	0.904

Table 27: Top 5 supported sequences for the fridge in household 2

FREEZER The freezer shows the same pattern as the fridge as reported in Table 28. Just as with the fridge the freezer seems to be used when the air exhaust is on but also after the fridge has been used. While the fridge had a support over 50% for almost all the supported sequences the freezer has a support over 65% for all the supported sequences. This might indicate that the freezer simply is used often, as it has to keep the temperature low in cycles, and therefore often shows up in sequences. This doesn't have to mean that the use of the freezer is user triggered.

Sequence	Occurrences	Support
Air exhaust ⇒ TV ⇒ Fridge ⇒ Freezer	195	0.965
Air exhaust ⇒ Fridge ⇒ TV ⇒ Freezer	206	0.963
Air exhaust ⇒ Fridge ⇒ Stereo ⇒ Freezer	320	0.955
Air exhaust ⇒ Fridge ⇒ Air exhaust ⇒ Freezer	353	0.954
Air exhaust ⇒ Fridge ⇒ Fridge ⇒ Freezer	379	0.953

Table 28: Top 5 supported sequences for the freezer in household 2

LAMP By intuition you would say that the lamp is used without a relation to any of the other devices, it must be related to the time of the day. The sequences in Table 29 support this theory as the most supported sequences

are made up by the lamp itself together with the fridge and freezer, which are on in cycles. Further the model shows that each sequence where the lamp is part of the support is over 94% while the sequences where the lamp is not part of the support is under 18%.

Sequence	Occurrences	Support
Lamp ⇒ Fridge ⇒ Freezer ⇒ Lamp	1166	0.984
Lamp ⇒ Fridge ⇒ Stereo ⇒ Lamp	2080	0.983
Lamp ⇒ Fridge ⇒ Fridge ⇒ Lamp	1997	0.983
Lamp ⇒ Stereo ⇒ Fridge ⇒ Lamp	2078	0.982
Lamp ⇒ Freezer ⇒ Fridge ⇒ Lamp	1161	0.982

Table 29: Top 5 supported sequences for the lamp in household 2

TV The TV is a device which, likely, will be used for longer periods in a row. For this reason the supported sequences should contain the TV and, once again, the fridge and freezer, in the sequences. This is approved by the sequences in Table 30. An interesting sequence in this table is that people tend to watch TV while the air exhaust is used. Data which is not in the table, but is in the model, indicates that the TV is mostly used when the lamp and stereo are on. This are sequences where the TV itself is not part of but the probability that the TV will be used in the near future is 92%. At the same time the probability is 87% when only the lamp is on and 72% when only the stereo is on.

Sequence	Occurrences	Support
TV ⇒ Fridge ⇒ Freezer ⇒ TV	9197	0.994
Stereo ⇒ Fridge ⇒ Freezer ⇒ TV	9271	0.994
TV ⇒ Freezer ⇒ Fridge ⇒ TV	9150	0.993
Stereo ⇒ Freezer ⇒ Fridge ⇒ TV	9227	0.993
Air exhaust ⇒ TV ⇒ Fridge ⇒ TV	519	0.993

Table 30: Top 5 supported sequences for the TV in household 2

STEREO The data shown in Table 31 indicates that the stereo always is used when the air exhaust is on. This is the only device in all the data sets that shows a 100% relationship between two devices. The model also tells that the stereo is on 94% of the time when only the lamp is on and almost 100% when the TV is used. Combining the results from the lamp, TV and stereo it can be concluded that the inhabitants of this household often watch TV with the stereo on at nights as the lamp has to be used.

Sequence	Occurrences	Support
Air exhaust ⇒ Freezer ⇒ Freezer ⇒ TV	302	1.0
Air exhaust ⇒ Freezer ⇒ Air exhaust ⇒ TV	234	1.0
Air exhaust ⇒ Freezer ⇒ Fridge ⇒ TV	181	1.0
Air exhaust ⇒ Freezer ⇒ TV ⇒ Freezer	290	1.0
Air exhaust ⇒ Freezer ⇒ TV ⇒ TV	310	1.0

Table 31: Top 5 supported sequences for the stereo in household 2

A.3.2 *Device specific validations*

In Table 32 the same data as in Table 4 is listed for convenience. In this section for each of the devices an attempt is made to explain the results by looking at the certainty of the predictions per device.

Statistic	Dishwasher	Air exhaust	Fridge	Freezer	Lamp	TV	Stereo
Correct	0,9929	0,9898	0,6662	0,6586	0,9852	0,8851	0,9728
Correct off	0,9935	0,9903	0,7733	0,6948	0,9854	0,9962	0,9914
Correct on	0,9364	0,7716	0,6053	0,651	0,9846	0,7024	0,943
Wrong	0,0072	0,0103	0,3339	0,3415	0,0149	0,115	0,0273
Device off	0,9842	0,9886	0,532	0,409	0,8247	0,732	0,6321
Device on	0,0159	0,0115	0,4681	0,5911	0,1754	0,2681	0,368
On coverage	0,5908	0,1475	0,8244	0,9106	0,9303	0,991	0,9855
Off coverage	0,9994	0,9995	0,527	0,2943	0,9969	0,8463	0,9653

Table 32: State prediction results for household 2

DISHWASHER The certainties of the dishwasher show a parabola like function in Figure 29 as expected. It is obvious that the dishwasher is barely ever used as the model is certain over 95% of the time that the dishwasher will not be used and achieves an accuracy of almost 100% with that. The little time that the dishwasher is used the model is capable of predicting this correctly when the certainty is 0.6 or higher, for 0.4 and 0.5 the model fails.

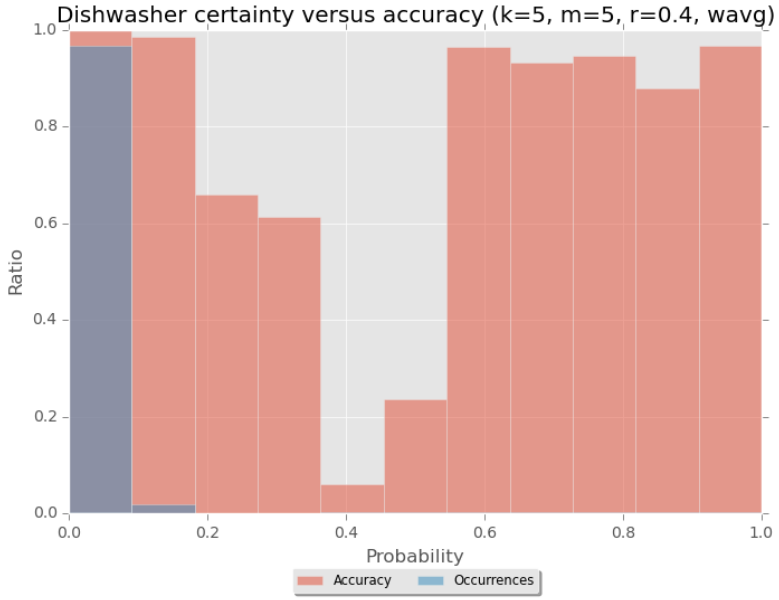

Figure 29: Histogram of certainty of predictions for the dishwasher in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

AIR EXHAUST The air exhaust shows an even stronger certainty that it will not be used. During 99% of the time the model is certain that the air exhaust absolutely will not be used with an accuracy of over 99%. At the same time the model is never really certain that the air exhaust will be used as there are no predictions with a certainty of 0.8 or higher. This observation conflicts with the supported sequences from Table 26 which indicate that there are multiple sequences with a probability over 80%. A possible explanation for this is that the weighted average, *wavg*, is used to predict and those sequences do not occur often and thus can be overwhelmed by sequences with a low probability but occur often (if a sequence with probability 0.2 occurs 1000 times and a sequence with 0.8 only 10 times the certainty will only be 0.2). This might mean that the air exhaust will achieve better predicting results when the average, *avg*, predictor method is used.

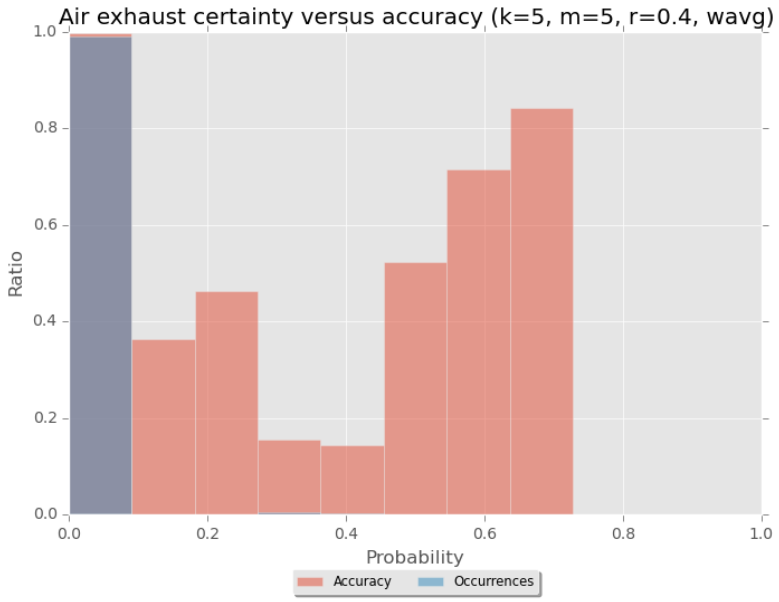

Figure 30: Histogram of certainty of predictions for the air exhaust in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

FRIDGE The data in Figure 31 shows that the fridge is a hard to predict device. During 40% of the time the model is uncertain whether it will be used or not and is achieving an accuracy of only 50%. The lack of predictions for the certainties 0.1-0.5, 0.9 and 1.0 support this observation. Overall the model is either really sure that the fridge will not be used, and still only achieving 80% accuracy, or the model is uncertain what will happen but tends to the fridge being used.

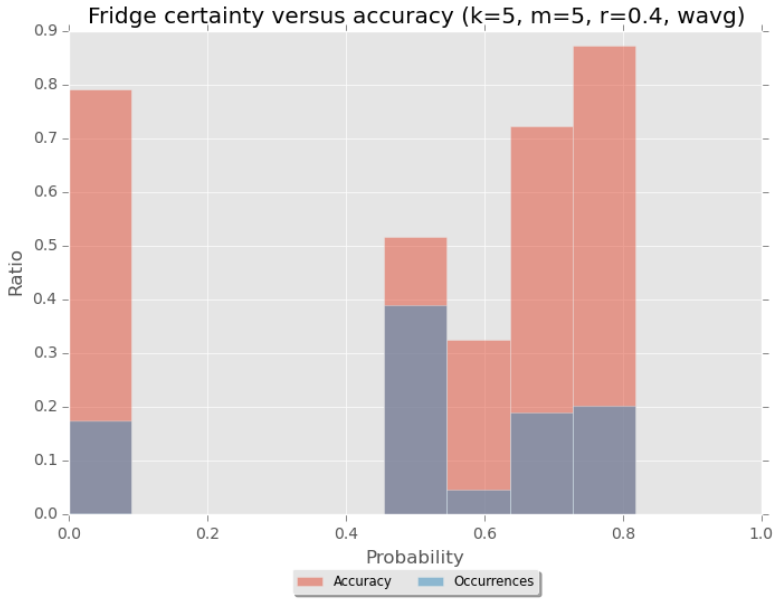

Figure 31: Histogram of certainty of predictions for the fridge in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

FREEZER The freezer shows a similar distribution as the fridge, this is likely due to the behaviour of those devices. Both the fridge and the freezer are turned on automatically in a cycle as it has to maintain the internal temperature. Just as the fridge the freezer also has a limited accuracy for predictions with the certainty of 0.6. While around 30% of the time the certainty of the model is 0.6 but it only achieves an accuracy of 30% while you expect an accuracy of, at least, above 50%. At the same time the accuracy for a certainty of 0.0 is limited for how confident the model is. An accuracy of 70% is way too low for a model that is absolutely certain that the device will not be used.

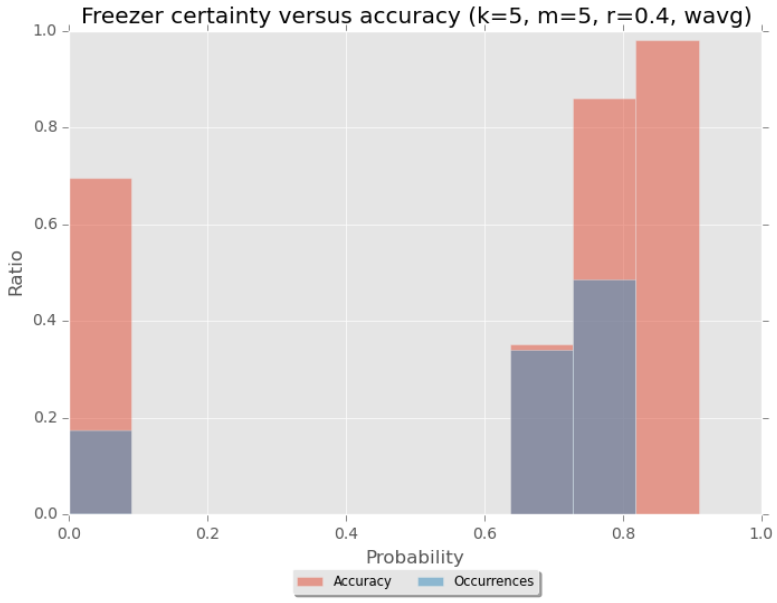

Figure 32: Histogram of certainty of predictions for the freezer in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

LAMP The lamp shows a regular pattern with a drop for the certainties 0.3 and 0.4. This can be explained by the minimal times those certainties occur, a few errors have much more impact on the accuracy than if it occurs thousands of times. Although 0.2 shows about the same number of occurrences (near 0%) this achieves in a better accuracy because the model is more certain about those predictions. A more surprising observation is that after a certainty of 0.6 the accuracy is, slowly, dropping from almost 100% to around 95%.

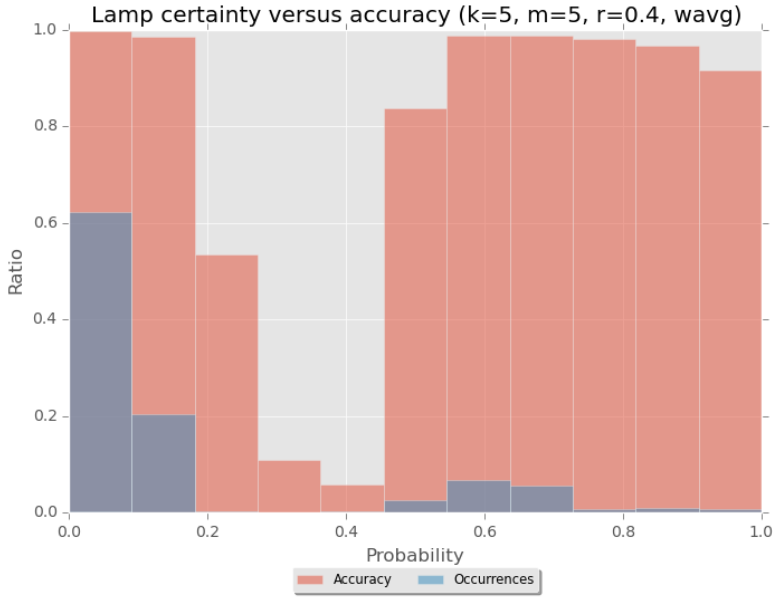

Figure 33: Histogram of certainty of predictions for the lamp in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

TV Figure 34 shows the certainty of predictions for the TV. The graph shows a rare phenomena where the accuracy with a certainty in the range 0.6-0.8 is really low while for all the other certainties the accuracy is great. The terrible results for the earlier mentioned certainties are hard to explain by the model, the only possible explanation is the transition from on to off or vice versa but this is hard to prove. *TODO: Look at the sequences who cause this terrible accuracy, has to be done in code*

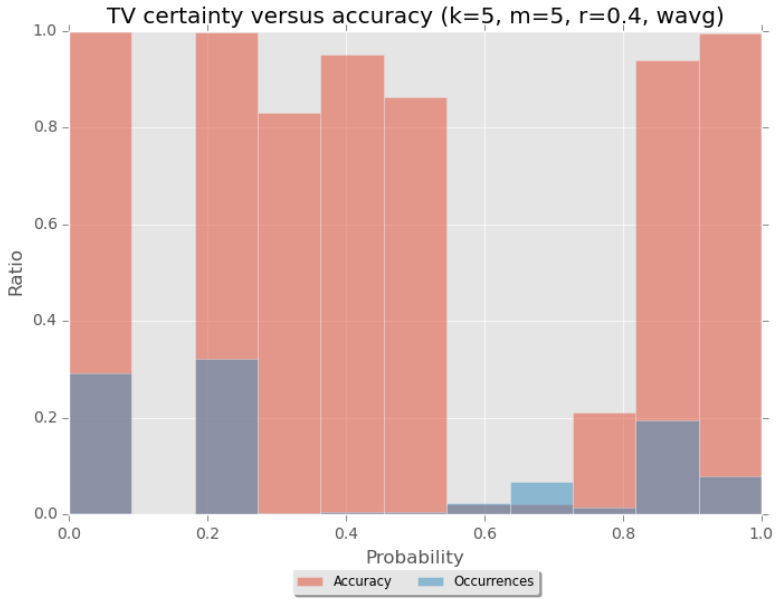

Figure 34: Histogram of certainty of predictions for the TV in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

STEREO Table 30 reported that there are sequences in which 100% of the time the stereo would be used in the near future. The data in Figure 35 supports this by achieving a near 100% accuracy for the stereo when the certainty is 1.0. Besides the accuracy this certainty occurs for about 25% of the measurements, which is by far the most for all devices measured. Overall the stereo predicts for 50% of the time a certainty of either 0.0 or 1.0 with an accuracy near 100% which indicates that the stereo is actually used in sequences and thus is good to predict. This is in line with the score for the stereo in Table 32 which reports an accuracy of 97%.

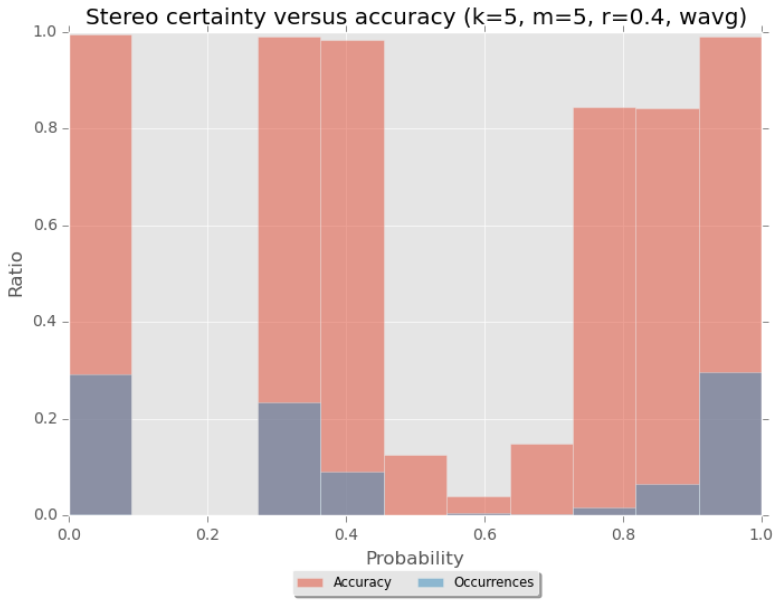

Figure 35: Histogram of certainty of predictions for the stereo in household 2 ($r = 0.4, k = 5, m = 5, wavg$)

B

DEVICE SPECIFIC ENERGY CONSUMPTION PATTERNS

Device specific consumption patterns can give better insight in the achieved results. For this reason an isolated plot is provided for each device.

B.1 RUG DATA SET

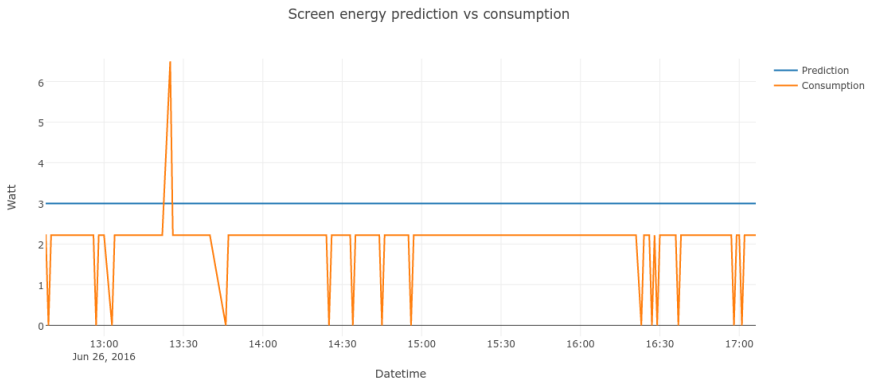

Figure 36: Energy consumption prediction for the screen in RUG data set

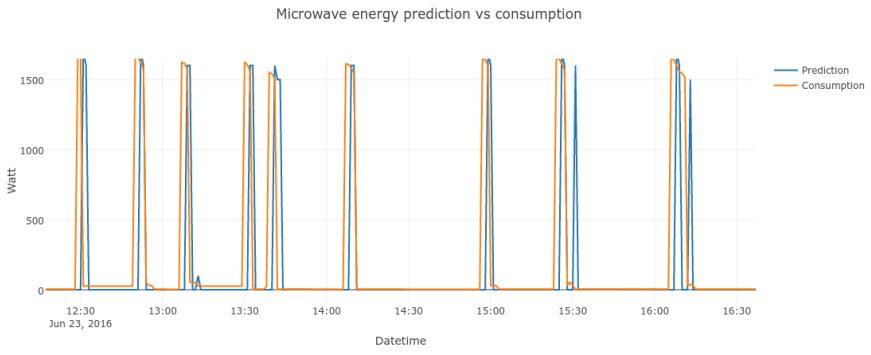

Figure 37: Energy consumption prediction for the microwave in RUG data set

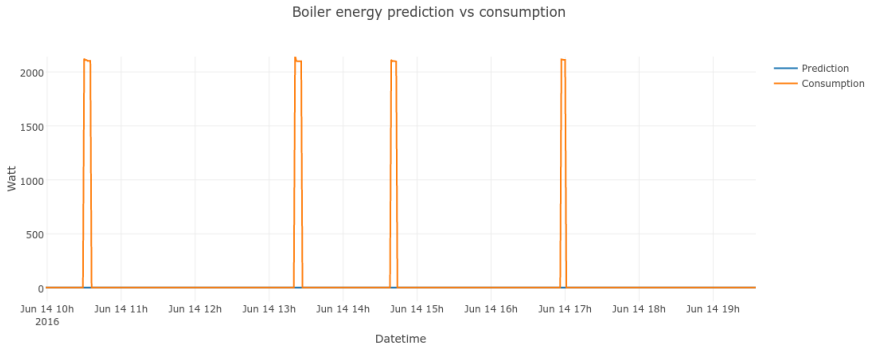

Figure 38: Energy consumption prediction for the boiler in RUG data set

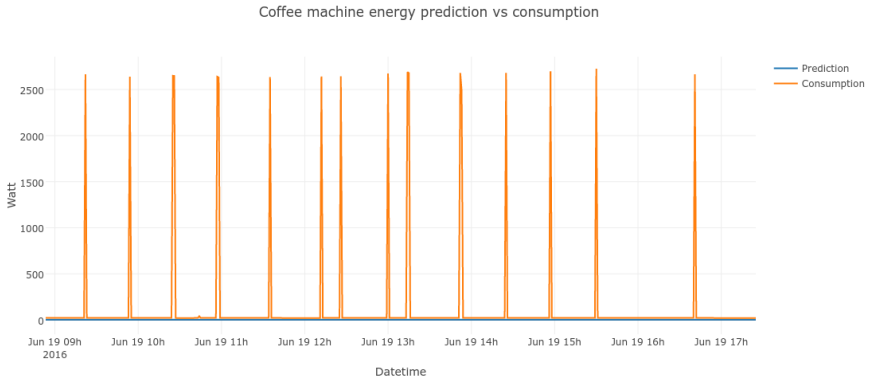

Figure 39: Energy consumption prediction for the coffee maker in RUG data set

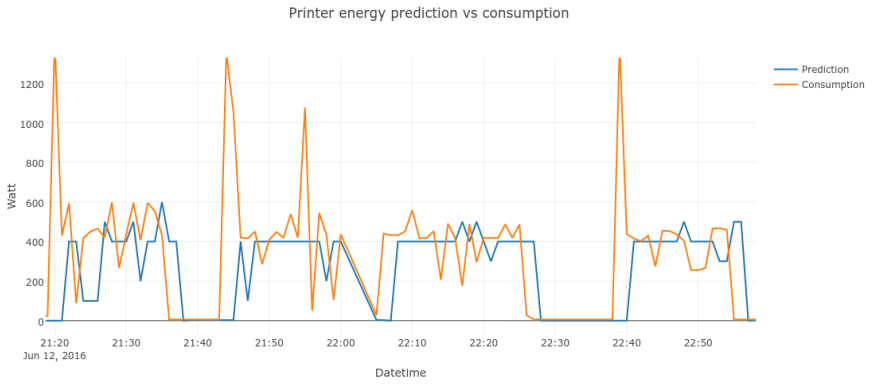

Figure 40: Energy consumption prediction for the printer in RUG data set

B.2 HOUSEHOLD 1

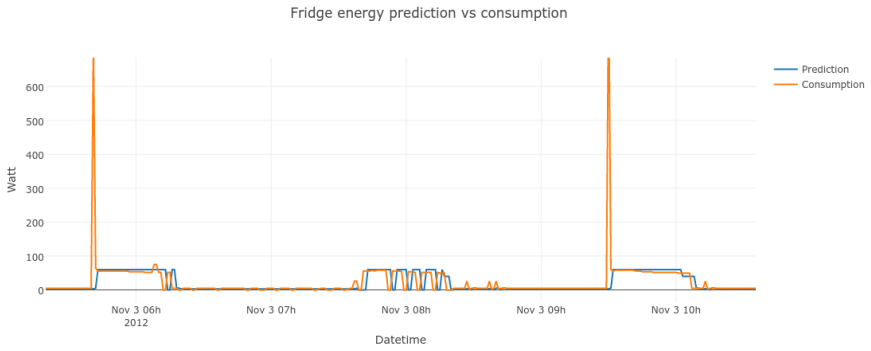

Figure 41: Energy consumption prediction for the fridge in household 1

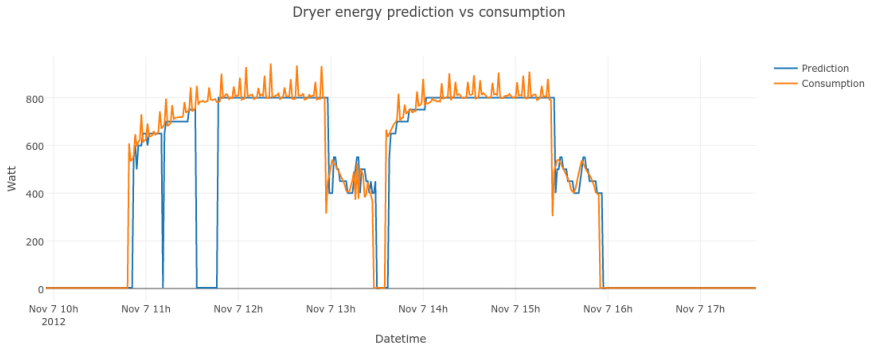

Figure 42: Energy consumption prediction for the dryer in household 1

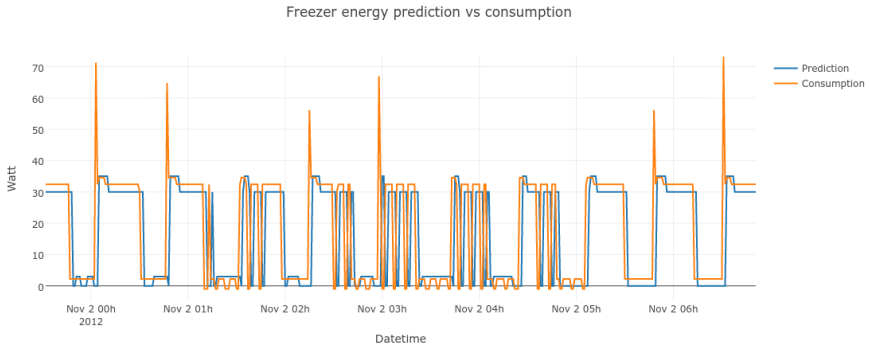

Figure 43: Energy consumption prediction for the freezer in household 1

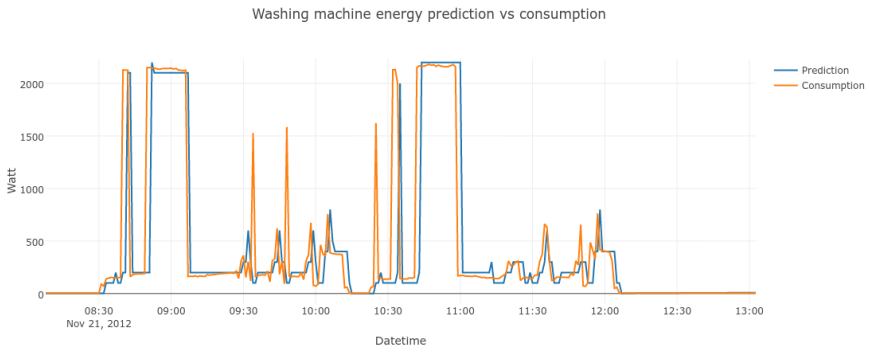

Figure 44: Energy consumption prediction for the washing machine in household 1

B.3 HOUSEHOLD 2

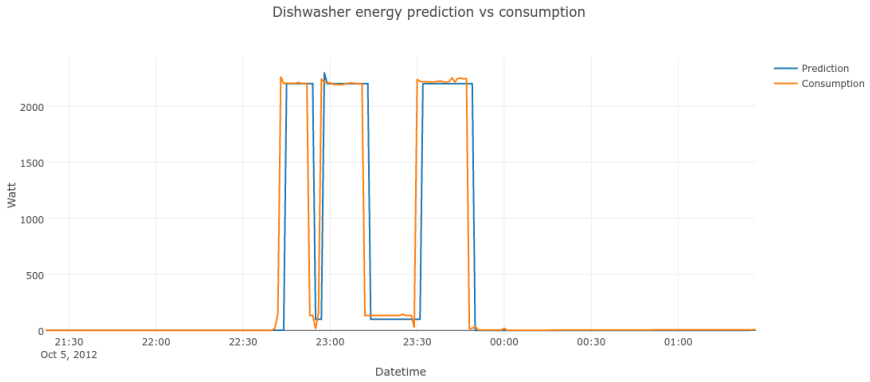

Figure 45: Energy consumption prediction for the dishwasher in household 2

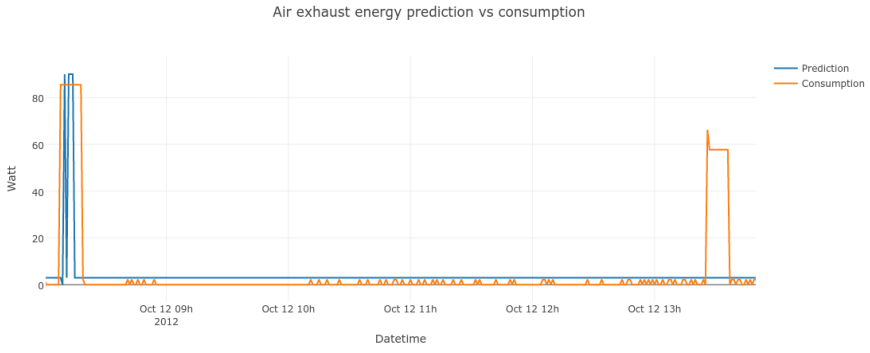

Figure 46: Energy consumption prediction for the air exhaust in household 2

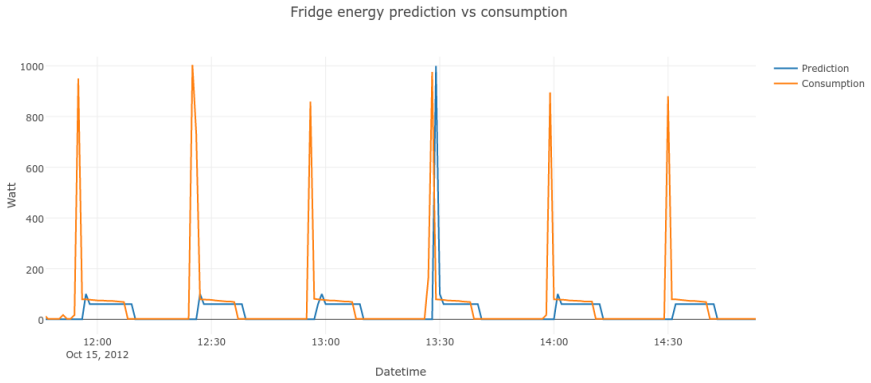

Figure 47: Energy consumption prediction for the fridge in household 2

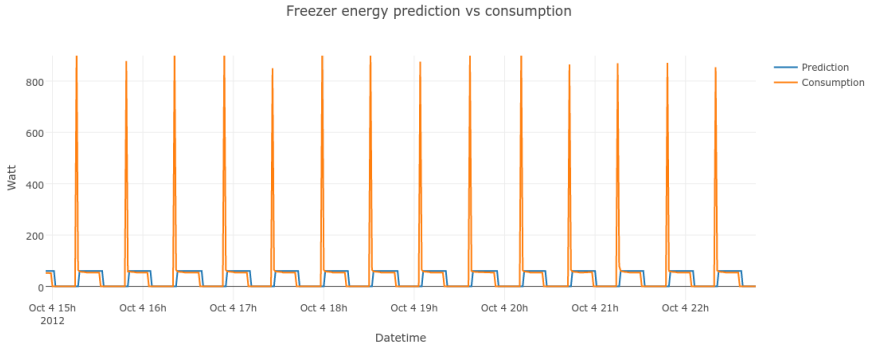

Figure 48: Energy consumption prediction for the freezer in household 2

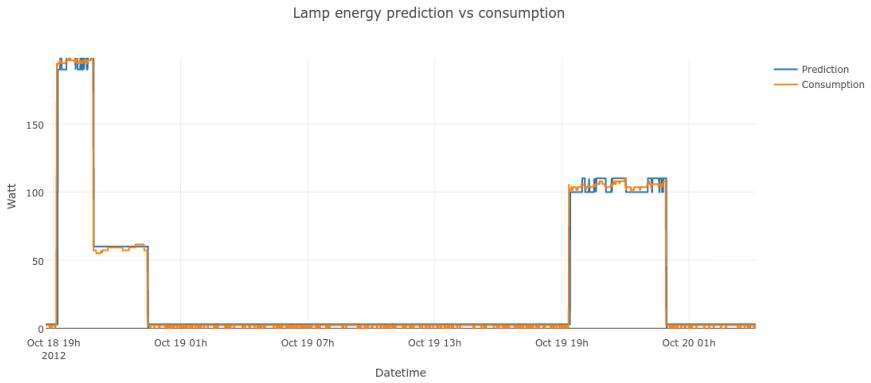

Figure 49: Energy consumption prediction for the lamp in household 2

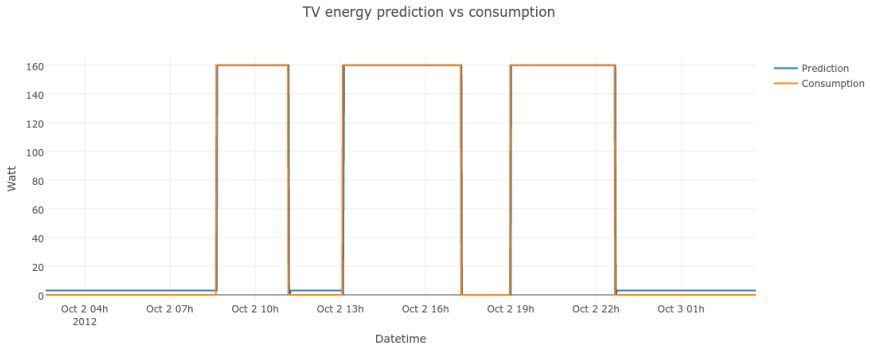

Figure 50: Energy consumption prediction for the TV in household 2

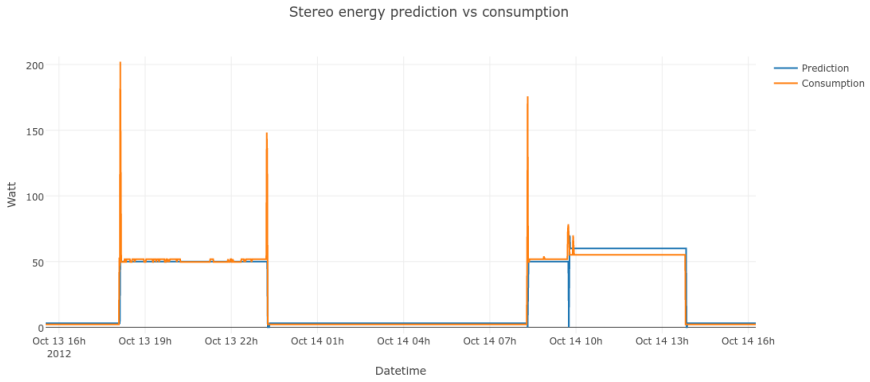

Figure 51: Energy consumption prediction for the stereo in household 2

C

DETAILED FAULT TOLERANCE RESULTS

Devices who have shown to be predictable while their own sensor was disabled are validated in this appendix. For this validation we will look at what sequences support the device and compare the results with the ground truth results, the normal state prediction. As there is not a single device in the RUG data set that is predictable this data set will not be validated in more detail.

C.1 HOUSEHOLD 1

For household 1 we look at each device which can be predicted and what sequences are supported and compare the results to the normal state prediction.

FRIDGE The fridge has 26 sequences where the probability is greater than 0.5 that the fridge will be on in the next 5 minutes, where has to be said that the highest probability is 0.511. This explains why the accuracy of the predictions is only around 60%, the probability is only a bit above 0.5. The sequences are listed in Table 33. The results for the fridge

Sequence	**Occurrences**	**Support**
Dryer ⇒ Freezer ⇒ Freezer ⇒ Dryer	1819	0.511
Dryer ⇒ Dryer ⇒ Freezer ⇒ Dryer	1890	0.511
Dryer ⇒ Freezer ⇒ Dryer ⇒ Dryer	1888	0.511
Dryer ⇒ Freezer ⇒ Dryer ⇒ Freezer	1753	0.510
Dryer ⇒ Dryer ⇒ Dryer ⇒ Freezer	1899	0.510

Table 33: Top 5 supported fault tolerance sequences for the fridge in household 1

The result for the fridge in Table 34 show significant less reliable results for the fault tolerance prediction than the ground truth. Where in the ground truth the fridge has an on coverage of 86% this is only 4,6% for the fault tolerance prediction. This can be explained with the sequences from Table 33 what tells us that the supported sequences only occur less than 2.000 times on a total of over 120.000 transactions.

Statistic	Fault tolerance	Ground truth
Correct	0,598	0,8764
Correct off	0,607	0,9082
Correct on	0,4071	0,8303
Wrong	0,4021	0,1237
Device off	0,6063	0,6063
Device on	0,3938	0,3938
On coverage	0,0463	0,8621
Off coverage	0,9563	0,8856
Same prediction	0,636	1

Table 34: Fault tolerance results for the fridge in household 1

FREEZER The freezer has with 34 sequences a few more supported sequences than the fridge. Those sequences also suggest a stronger relation with probabilities varying from 0.680 to 0.699 as seen in Table 35. As a result for around 67% of the prediction the prediction equals the ground truth

prediction as is listed in Table 36. Interesting is to see that the freezer has an on coverage of 45% which is a lot better than the fridge that only had 4%.

Sequence	Occurrences	Support
Dryer ⇒ Dryer ⇒ Dryer ⇒ Fridge	1914	0.699
Dryer ⇒ Dryer ⇒ Fridge ⇒ Fridge	1871	0.699
Dryer ⇒ Fridge ⇒ Dryer ⇒ Fridge	1829	0.698
Dryer ⇒ Dryer ⇒ Fridge	2068	0.698
Dryer ⇒ Dryer ⇒ Fridge ⇒ Dryer	3259	0.697

Table 35: Top 5 supported fault tolerance sequences for the freezer in household 1

Statistic	Fault tolerance	Ground truth
Correct	0,4919	0,7173
Correct off	0,3835	0,7198
Correct on	0,623	0,7167
Wrong	0,5082	0,2828
Device off	0,3806	0,3806
Device on	0,6195	0,6195
On coverage	0,4552	0,8993
Off coverage	0,5516	0,4213
Same as ground truth	0,6754	1

Table 36: Fault tolerance results for the freezer in household 1

C.2 HOUSEHOLD 2

The fridge, freezer, TV and stereo are devices which can be predicted in a fault tolerance situation for household 2. An interesting observation is that the most of supported sequences are based on the air exhaust which is barely ever used.

FRIDGE There are 328 sequences with support for the fridge without the presence of the fridge itself. The five best supported sequences from Table 37 show that the air exhaust is related to the fridge. Although the sequences are rare they still show that the fridge will be used after it for 70% of the time.

Sequence	Occurrences	Support
Air exhaust ⇒ Air exhaust ⇒ Air exhaust ⇒ Air exhaust	524	0.718
Air exhaust ⇒ Air exhaust ⇒ TV ⇒ Air exhaust	225	0.714
Air exhaust ⇒ Stereo ⇒ TV ⇒ Air exhaust	225	0.713
Air exhaust ⇒ TV ⇒ Stereo ⇒ Air exhaust	225	0.713
Air exhaust ⇒ TV ⇒ Air exhaust ⇒ Air exhaust	224	0.713

Table 37: Top 5 supported fault tolerance sequences for the fridge in household 2

Comparing the results for the fridge to the ground truth in Table 38 shows that the overall accuracy has dropped for about 12% which is not that bad. The predictions are the same as before for 75% of the time which causes the on coverage to drop from 82% to 42%. Overall the fridge was not a device which was easy predictable in the normal state prediction so this drop in accuracy is expected and not extraordinary.

FREEZER The freezer shows a similar pattern as the fridge. The table of best supported sequences, listed in Table 39, is dominated by sequences starting with the air exhaust. Just as with the fridge the best supported

Statistic	Fault tolerance	Ground truth
Correct	0,544	0,6662
Correct off	0,5619	0,7733
Correct on	0,5156	0,6053
Wrong	0,4561	0,3339
Device off	0,532	0,532
Device on	0,4681	0,4681
On coverage	0,4248	0,8244
Off coverage	0,6489	0,527
Same prediction	0,7482	1

Table 38: Fault tolerance results for the fridge in household 2

sequences are rare but the probability that the freezer will be used for this sequences is as high as 86%.

Sequence	**Occurrences**	**Support**
Air exhaust ⇒ Fridge ⇒ TV ⇒ Stereo	219	0.864
Air exhaust ⇒ Fridge ⇒ TV ⇒ TV	217	0.863
Air exhaust ⇒ Fridge ⇒ TV ⇒ Fridge	180	0.861
Air exhaust ⇒ Fridge ⇒ Fridge ⇒ TV	196	0.860
Air exhaust ⇒ Fridge ⇒ TV ⇒ Air exhaust	163	0.859

Table 39: Top 5 supported fault tolerance sequences for the freezer in household 2

During 81% of the time the prediction for the freezer equals the normal state prediction experiment. Table 40 reports a total drop of 10% for the accuracy of the predictions. An interesting statistic is that the accuracy of predictions that the device will be used almost stays the same, it only drops 4%.

Statistic	Fault tolerance	Ground truth
Correct	0,5518	0,6586
Correct off	0,4458	0,6948
Correct on	0,612	0,651
Wrong	0,4483	0,3415
Device off	0,409	0,409
Device on	0,5911	0,5911
On coverage	0,6601	0,9106
Off coverage	0,3952	0,2943
Same as ground truth	0,8107	1

Table 40: Fault tolerance results for the freezer in household 2

TV The best supported sequences in Table 41 show a really high support when the lamp is on. The sequences indicate that the inhabitants watch TV at night because the lamp is on.

Table 42 shows great results when you compare the fault tolerance predictions to the ground truth. The TV is predicted the same for 99,98% of the time with his own sensor disabled as when it is enabled. As a result the accuracy of the predictions is differs only 0,01%. This proofs that it really is possible to predict a device state as long there is a strong correlation between one or more other devices.

STEREO According to Table 43 the inhabitants of household 2 always have the stereo on while they are using the air exhaust and the TV.

The stereo shows some great results just as the TV. The stereo is predicted the same for 92% of the time with a total accuracy of 89% which is 8% lower

Sequence	**Occurrences**	**Support**
Lamp ⇒ Freezer ⇒ Lamp ⇒ Stereo	2754	0.921
Lamp ⇒ Stereo ⇒ Lamp ⇒ Stereo	4767	0.921
Lamp ⇒ Lamp ⇒ Lamp ⇒ Stereo	4776	0.921
Lamp ⇒ Freezer ⇒ Stereo ⇒ Lamp	2734	0.920
Lamp ⇒ Lamp ⇒ Stereo ⇒ Stereo	4795	0.920

Table 41: Top 5 supported fault tolerance sequences for the TV in household 2

Statistic	Fault tolerance	Ground truth
Correct	0,885	0,8851
Correct off	0,9959	0,9962
Correct on	0,7025	0,7024
Wrong	0,1151	0,115
Device off	0,732	0,732
Device on	0,2681	0,2681
On coverage	0,9905	0,991
Off coverage	0,8464	0,8463
Same as ground truth	0,9998	1

Table 42: Fault tolerance results for the TV in household 2

than the ground truth. The difference in accuracy is mostly caused by the on coverage which dropped from 99% to 77% as can be seen in Table 44.

Sequence	**Occurrences**	**Support**
Air exhaust ⇒ TV	510	1.0
Air exhaust ⇒ TV ⇒ TV	498	1.0
Air exhaust ⇒ TV ⇒ TV ⇒ TV	485	1.0
Air exhaust ⇒ Air exhaust ⇒ TV	455	1.0
Air exhaust ⇒ Air exhaust ⇒ TV ⇒ TV	443	1.0

Table 43: Top 5 supported fault tolerance sequences for the stereo in household 2

Statistic	Fault tolerance	Ground truth
Correct	0,8932	0,9728
Correct off	0,8763	0,9914
Correct on	0,932	0,943
Wrong	0,1069	0,0273
Device off	0,6321	0,6321
Device on	0,368	0,368
On coverage	0,7654	0,9855
Off coverage	0,9675	0,9653
Same as ground truth	0,9177	1

Table 44: Fault tolerance results for the stereo in household 2

[1] Dsg - research project: Smart meter services. http://www.vs.inf.ethz.ch/res/show.html?what=smart-meter-services. Accessed: 30-09-2016.

[2] Dutch parliament votes to close down country's coal industry. https://www.theguardian.com/environment/2016/sep/23/dutch-parliament-votes-to-close-down-countrys-coal-industry. Accessed: 25-10-2016.

[3] Energy management systems of the 21st century | plugwise. http://www.plugwise.com. Accessed: 30-09-2016.

[4] Heating and cooling no longer majority of u.s. home energy use. https://www.eia.gov/todayinenergy/detail.php?id=10271. Accessed: 25-10-2016.

[5] Landmark court ruling tells dutch government to do more on climate change. http://www.nature.com/news/landmark-court-ruling-tells-dutch-government-to-do-more-on-climate-change-1.17841. Accessed: 25-10-2016.

[6] Residential monitoring to decrease energy use and carbon emissions in europe. http://remodece.isr.uc.pt. Accessed: 24-10-2016.

[7] R Agrawal, T Imielinski, and A Swami. Mining associations between sets of items in massive databases. In *Proceedings of the ACM-SIGMOD Int'l Conference on Management of Data*, pages 207–216, 1993.

[8] Christina Athanasopoulou and Vasilis Chatziathanasiou. Intelligent system for identification and replacement of faulty sensor measurements in Thermal Power Plants (IPPAMAS: Part 1). *Expert Systems with Applications*, 36(5):8750–8757, 2009.

[9] Asier Aztiria, Juan Carlos, and Diane J Cook. Discovering frequent user âĂŞ environment interactions in intelligent environments. pages 91–103, 2012.

[10] Asier Aztiria, Alberto Izaguirre, Rosa Basagoiti, Juan Carlos Augusto, and Diane J. Cook. Automatic modeling of frequent user behaviours in intelligent environments. *Proceedings - 2010 6th International Conference on Intelligent Environments, IE 2010*, pages 7–12, 2010.

[11] A. Barbato, A. Capone, M. Rodolfi, and D. Tagliaferri. Forecasting the usage of household appliances through power meter sensors for demand management in the smart grid. *2011 IEEE International Conference on Smart Grid Communications, SmartGridComm 2011*, pages 404–409, 2011.

[12] Kaustav Basu, Lamis Hawarah, Nicoleta Arghira, Hussein Joumaa, and Stephane Ploix. A prediction system for home appliance usage. *Energy and Buildings*, 67:668–679, 2013.

[13] Christian Beckel, Wilhelm Kleiminger, Romano Cicchetti, Thorsten Staake, and Silvia Santini. The eco data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM International Conference on Embedded Systems for Energy-Efficient Buildings (BuildSys 2014). Memphis, TN, USA*, pages 80–89. ACM, November 2014.

[14] Abdullatif E. Ben-Nakhi and Mohamed A. Mahmoud. Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management*, 45(13-14):2127–2141, 2004.

[15] CBS. Elektriciteit in Nederland. pages 19–21, 2015.

[16] M Çunkaş and AA Altun. Long term electricity demand forecasting in turkey using artificial neural networks. *Energy Sources, Part B: Economics, Planning, and Policy*, 5(3):279–289, 2010.

[17] Viktoriya Degeler, Alexander Lazovik, Francesco Leotta, and Massimo Mecella. Itemset-based mining of constraints for enacting smart environments. *2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014*, pages 41–46, 2014.

[18] M Deshpande and G Karypis. Selective Markov Models for Predicting Web-Page Accesses. *ACM Transactions on Internet Technology*, 4(2):163–184, 2004.

[19] Michelangelo Diligenti, Paolo Frasconi, and Marco Gori. Hidden tree markov models for document image classification. *IEEE Transactions on pattern analysis and machine intelligence*, 25(4):519–523, 2003.

[20] Ricardo Dunia, Fisher-rosemount Systems, S Joe Qin, Thomas F Edgar, and Thomas J Mcavoy. Identification of Faulty Sensors Using Principal Component Analysis. *Process Systems Engineering*, 42(10):2707–2812, 1996.

[21] Thi V Duong, Hung Hai Bui, Dinh Q Phung, and Svetha Venkatesh. Activity recognition and abnormality detection with the switching hidden semi-markov model. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 838–845. IEEE, 2005.

[22] Richard E. Edwards, Joshua New, and Lynne E. Parker. Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings*, 49:591–603, 2012.

[23] Magdalini Eirinaki, Michalis Vazirgiannis, and Dimitris Kapogiannis. Web path recommendations based on page ranking and markov models. In *Proceedings of the 7th annual ACM international workshop on Web information and data management*, pages 2–9. ACM, 2005.

[24] Paul M Frank. Fault diagnosis in dynamic systems using analytical knowledge-based redundancy âĂŞ A survey and some new results. *Automatica*, 26(3):459–474, 1990.

[25] MDKZ Ghiassi, David K Zimbra, and H Saidane. Medium term system load forecasting with a dynamic artificial neural network model. *Electric Power Systems Research*, 76(5):302–316, 2006.

[26] T H Guo and J Nurre. Sensor failure detection and recovery by neural networks. *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, 1:221–226, 1991.

[27] Pablo H. Ibargüengoytia, Sunil Vadera, and L. Enrique Sucar. A probabilistic model for information and sensor validation. *Computer Journal*, 49(1):113–126, 2006.

[28] Ian Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[29] Radiša Jovanović, Aleksandra A. Sretenović, and Branislav D. Živković. Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, 94:189–199, 2015.

[30] Hyun Chul Jung, Jin Sung Kim, and Hoon Heo. Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. *Energy and Buildings*, 90:76–84, 2015.

[31] Eleftheria Katsiri. Principles of context inferences. In *ADJUNCT PROCEEDINGS*, page 33, 2002.

[32] Gaetan Kerschen, Pascal De Boe, Jean-claude Golinval, and Keith Worden. Sensor validation using principal component analysis. *Smart Materials and Structures*, 14(1):36–42, 2004.

[33] Wilhelm Kleiminger, Christian Beckel, and Silvia Santini. Household occupancy monitoring using electricity meters. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, Osaka, Japan, September 2015.

[34] KY Lee, YT Cha, and JH Park. Short-term load forecasting using an artificial neural network. *IEEE Transactions on Power Systems*, 7(1):124–132, 1992.

[35] Kangji Li, Chenglei Hu, Guohai Liu, and Wenping Xue. Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis. *Energy and Buildings*, 108:106–113, 2015.

[36] Qiong Li, Qinglin Meng, Jiejin Cai, Hiroshi Yoshino, and Akashi Mochida. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*, 50(1):90–96, 2009.

[37] Sebastian Lühr, Geoff West, and Svetha Venkatesh. Recognition of emergent human behaviour in a smart home: A data mining approach. *Pervasive and Mobile Computing*, 3(2):95–116, 2007.

[38] Nasir Mehranbod, Masoud Soroush, and Chanin Panjapornpon. A method of sensor fault detection and identification. *Journal of Process Control*, 15(3):321–339, 2005.

[39] Ara V Nefian and Monson H Hayes. Face detection and recognition using hidden markov models. In *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, volume 1, pages 141–145. IEEE, 1998.

[40] Ecole Centrale Paris, Applied Mathematics, and Voie Vignes. [hal-00617930, v1] Parallel Support Vector Machines Applied to the Prediction of Multiple Buildings Energy Consumption_[hal-00617930, v1] Parallel Support Vector Machines Applied to the Prediction of Multiple Buildings Energy Consumption_8.pdf. 4(2):1–17, 2010.

[41] James Pitkow and Peter Pirolli. Mining longest repeating subsequences to predict world wide web surfing. *Proceedings of USITS '99: The 2nd USENIX Symposium on Internet Technologies \& Systems*, pages 139–150, 1999.

[42] Parisa Rashidi, Diane J Cook, Lawrence B Holder, and Maureen Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. *IEEE transactions on knowledge and data engineering*, 23(4):527–539, 2011.

[43] Ihab Samy, Ian Postlethwaite, and Da Wei Gu. Survey and application of sensor fault detection and isolation schemes. *Control Engineering Practice*, 19(7):658–674, 2011.

[44] Johan AK Suykens. Nonlinear modelling and support vector machines. In *Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE*, volume 1, pages 287–294. IEEE, 2001.

[45] Geoffrey K F Tso and Kelvin K W Yau. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, 32(9):1761–1768, 2007.

[46] Shengwei Wang and Youming Chen. Sensor validation and reconstruction for building central chilling systems based on principal component analysis. *Energy Conversion and Management*, 45(5):673–695, 2004.

[47] Zhaoxia Wang and Yan Ding. An occupant-based energy consumption prediction model for office equipment. *Energy and Buildings*, 109:12–22, 2015.

[48] James A Whittaker and Michael G Thomason. A markov chain model for statistical software testing. *IEEE Transactions on Software engineering*, 20(10):812–824, 1994.

[49] B M Wise and N L Ricker. Recent advances in nultivariate statistical process control improving robustness and sensitivity. *IFAC Symposium on Advanced Control of Chemical Processes, Toulouse, France, October*, page 17, 1991.

[50] Fan Zhang, Chirag Deb, Siew Eang Lee, Junjing Yang, and Kwok Wei Shah. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Buildings*, 126:94–103, 2016.

[51] Dmytro Zubov. An Energy-Saving Concept of the Smart Building Power Grid with Separated Lines for Standby Devices. (2014):148–156, 2016.