Imperial College London
Department Of Computing

Inferring Appliance-by-Appliance Energy Consumption from Whole-House Electricity Meter Readings

by

Konstantinos Kamnitsas [kk2412]

Submitted in partial fulfilment of the requirements for the
MSc Degree in Computing Science of Imperial College London

September 2013

Abstract

Energy conservation has been in the forefront of attention for quite some time. The climate change, the continuously increasing energy demands and the limited availability of energy resources have made several bodies look for means to achieve more efficient energy consumption. National targets have been set in order to limit the environmental effects of fossil fuel consumption. The residential sector accounts for a big part of the global energy consumption (approximately 26% for the UK). In order to monitor the national energy consumption, as well as motivate home owners to improve their electricity consumption behaviour, smart-meters are being deployed in a large scale in the UK and the USA.

A smart-meter is connected to the electric main of a residence and monitors the energy consumption of the whole house. The collected data are of great value; it may allow for prediction of the national load and enable optimization of the power grid or enable the utility companies to provide better services to their clients. Of great interest is the extraction of detailed information from this data regarding the electricity usage of consumers. Electricity is at the moment an 'invisible' product, in the sense that it is rather hard to estimate how much is consumed and by which appliance. Providing more detailed information than simply a monthly bill to the consumers would enable them to get control over their electricity consumption and make wiser use of it.

The task of extracting end-use and appliance-level information from the measurements of the whole-house consumption, the aggregate signal, is referred to as energy disaggregation. It has been the target of research attempts for over 20 years. However, there has not been developed a robust and accurate technique that could be widely adopted in the residential sector yet. The developed systems either require special equipment or have important limitations.

This research elaborates on the disaggregation of multi-state appliances that are commonly found in residences. Multi-state appliances are problematic to disaggregate using classic approaches, which are based only on the detection of steady states in the power signal, because of the complicated nature of their operation.

A novel supervised pattern recognition technique has been developed, inspired by feature extraction techniques used in Optical Character Recognition. The designed system attempts to determine the 'approximate shape' of the signatures (the measured power consumption of the appliance during one single operation) of a multi-state appliance by detecting the main parts that they consist of. By processing the signatures of various models of the same appliance type, the system attempts to extract patterns and features that are frequently observed during the operation of an appliance. The extracted set of features comprises the *profile* of an appliance, which describes the generic shape of a multi-state appliance's signatures. It is an attempt towards the creation of a generic signature of an appliance, that may describe the appliance type as a whole, independent of the model's size, manufacturer or other such factors.

Most previous research attempts performed on data sampled at the frequency range of 1Hz-0.1Hz only extract the power consumed during steady states of an appliance's operation. In this work various features have been investigated, in order to determine a richer set that can accurately describe an appliance, allowing its detection and at the same time distinguish it from the rest. The investigated features include durations of states, time intervals between different states, repeating parts in an appliance's signature, periods of rapid changes created by an appliance's operation, power changes with the form of spikes and more. Signatures of different models of appliances have been processed, available in various public datasets. Three multi-state appliances were the main targets of this work: the dishwasher, the washing machine and the dryer.

A system has been implemented in order to extract the various features, generate the profile of each appliance and, finally, test the approach by disaggregating operations of these appliances. The profiles generated manage to describe the generic characteristics of an appliance, allowing the system to detect fairly accurate instances of all three appliances in an aggregate signal (50-100% disaggregation accuracy in the tested cases). The results were particularly promising taking into account the very low number of false positives (83-100% detection accuracy in the tested cases), fact that may allow for further extensions to this work.

Acknowledgements

At this point I would like to thank my supervisor, Professor William Knottenbelt, as well as the co-supervisor of this project, PhD student Jack Kelly. Their deep passion and enthusiasm for their work has been a real inspiration since the beginning of this MSc course. Thank you for trusting me with this project, as well as your constant support and invaluable advice whenever it was needed.

But most importantly, from the bottom of my heart, I would like to thank my family for their limitless love and support throughout all these years of my life. During the good and during the tough times. This work is dedicated to them, my father, my mother and my little brother.

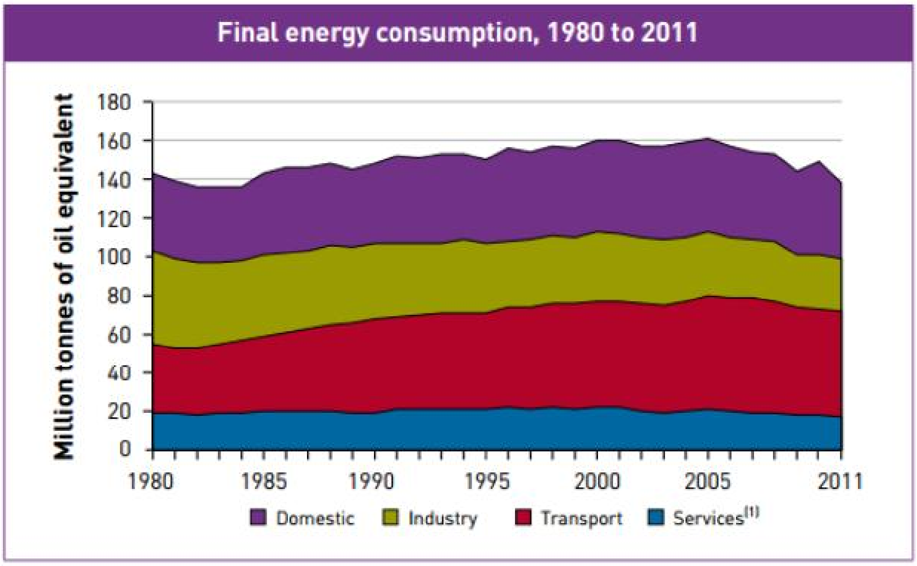
Father, Mother, Apostoli, thank you for everything.

Contents

1 Introduction

1.1 Motivation

Between the frequent discussions on the climate change and the annual increase of a households expenses on electricity bills, the effects of energy overconsumption have become pretty obvious to anyone nowadays. The energy problem has attracted the attention of many interested bodies. National targets have been set, in order for the global carbon emissions to be reduced. The research community is tackling different aspects of the problem, with much focus on optimizing the energy usage. An important part of the global energy consumption is due to the domestic sector. For instance, in the United Kingdom for the year 2011, approximately 26% of the national energy consumption was due to domestic consumption [1]. The biggest percentage of residential consumption was for space heating (60%) and water heating (18%). Electrical appliances and lighting account for approximately 22% of the energy consumed by households in the UK for the year 2011 [3].



2011 — **Million tonnes of oil equivalent**

	Industry	Domestic	Transport	Services¹	Total
Coal & manufactured fuels	1.7	0.8	0.0	0.0	2.5
Gas	10.7	25.2	-	6.5	42.4
Oil	4.5	2.7	53.7	1.4	62.3
Electricity	8.8	9.6	0.4	8.6	27.3
Bioenergy and heat	1.4	0.6	1.1	0.7	3.8
Total	**27.1**	**38.8**	**55.2**	**17.2**	**138.3**

(1) Includes agriculture

Figure 1.1: Energy consumption for the UK, per sector. The noticeable decrease in the consumption for the domestic sector between the years 2006 and 2011 was mainly caused by the relatively mild winters. Source:[1]

Furthermore, in the USA, electrical appliances and lighting were found to account approximately for the 35% of total domestic consumption of 2009 [8]. As presented in the latter research, electricity consumption has increased by 23% over the past decade in the USA, mostly because of increased use of air-conditioning and the introduction of new appliances. Projections show that electricity consumption in the residential vector will continue to increase, estimating consumption 20% higher in 2030 in comparison to 2007.

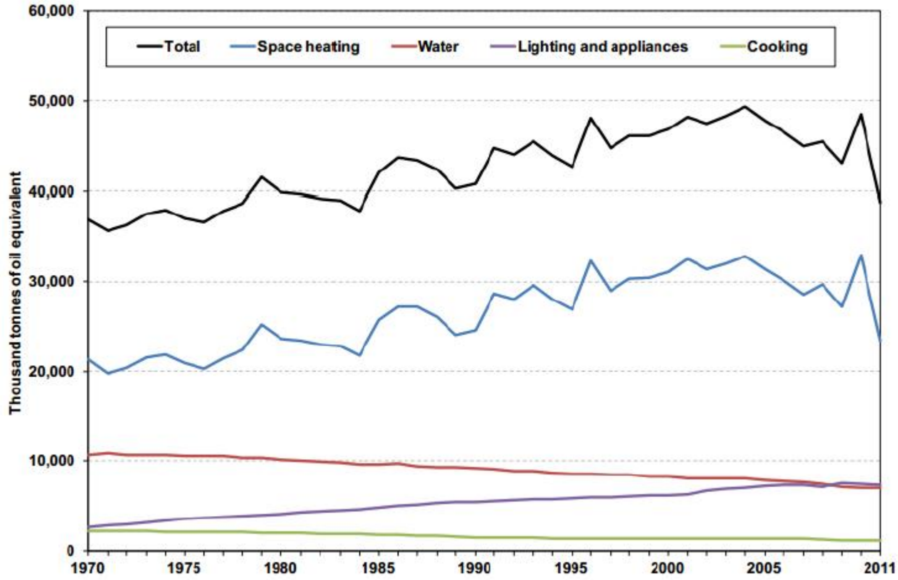


Figure 1.2: Domestic energy consumption for the UK, by the end use. Source:[3]

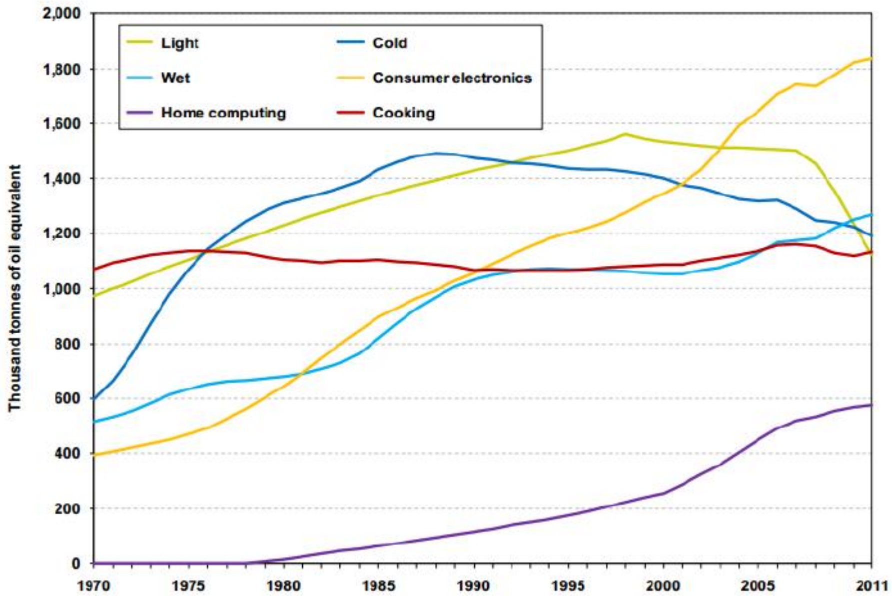


Figure 1.3: Electricity consumption by household domestic appliances, by broad type, UK, 1970 to 2011. Source:[3]

With the continuous introduction of new electrical appliances, computing equipment and gadgets, the need for electricity increases, pushing the electricity price higher and higher with every passing year. Only in the past decade, the price of electricity has been increased by approximately 50% in the UK [4].

The use of a fairly simple device, the smart meter, allows for more efficient use of energy in the domestic sector. A smart meter is installed on the electric main of the house and measures the total energy consumption of the house in regular intervals. The data can be of great use to many different bodies. For instance, utilities are enabled to provide better services to their clients and consumers may have more control over their electricity consumption. Actually, they tend to modify their behaviour towards more efficient use of electricity after they are presented with that kind of data [11], saving money and reducing their carbon footprint. Smart meters also enable the development of the "smart grid" by providing data that can be used to achieve, for instance, more efficient load - balancing. Smart meters will become even more important for the grid operation with the increase in micro-generation, for example photovoltaic panels, or the future increase in the use of electric cars.

Millions of smart-meters are already installed globally and their penetration is expected to rapidly increase [7]. The UK government is overseeing an ambitious programme in order to reduce the national carbon emissions and also help the consumers deal with the rising electricity prices. Energy suppliers will be required to install smart-meters in millions of homes and businesses, in a massive rollout that was scheduled to begin in 2014 and complete in 2019; On May 2013 the program was delayed for one year to achieve better planning. Some suppliers have already started installing smart-meters [5, 6].

The data gathered from the smart meters can be the source of various information. From detecting a malfunctioning device in the house which consumes unusual amounts of energy, to monitoring the use of electricity for heating or cooling so that the grid operator may predict the grid's load for the next winter or summer. This paper elaborates on the disaggregation of the data gathered for a whole house's consumption; the inference of the consumption of the individual appliances within the house. Research in this field has lasted well over 20 years, however still there seems to be no technique that can be widely adopted in the domestic sector without limitations. Previous research attempts have been quite successful in disaggregating certain common appliances with the use of custom equipment sampling the aggregate signal at high frequencies and detecting features, like transients, which are specific to those appliances. The use of special equipment makes these techniques rather unattractive for wide adoption in the domestic sector. Advancement in disaggregation at lower sampling rates has been quite slower, since less features can be detected, with each technique developed having each own limitations.

1.2 Objectives

Main aims of this research were:

- The development of a technique in order to form a generic profile of an appliance type. It is an attempt towards the creation of generic signatures for appliances, that would characterize an appliance type as a whole and enable the disaggregation of the appliance regardless the size of the model or the operating conditions.
- The investigation for features and patterns in the signatures of appliances, in order to determine a set of features that may allow for accurate detection of an appliance in the aggregate signal.
- The development of a disaggregation algorithm that, given a generated profile for an appliance type, detects with accuracy the operation of the appliance in the whole-house consumption signal.
- The developed methodology should be able to perform on measurements sampled at rates between 0.1Hz and 1Hz, in order to be applicable on data gathered from common smart-meters.

1.3 Contributions

This thesis presents a novel methodology of modelling an appliance with a generalized profile and disaggregating it by evaluating the existence or absence of a set of features.

The system designed 'maps' an appliance after processing the signatures (measured consumption of an appliance during one operation) provided for an appliance. This is achieved by detecting the main parts of an appliance's operation, the parts which are consistently observed in the signatures of different models. These main parts describe the 'rough shape' of the appliance's signature and are used to form a generic profile of the appliance type.

Most cases of previous related work on low resolution data extract and make use only of the power consumed in the various steady states of an appliance. This research differs radically from these approaches. Several power and non-power related features are investigated and extracted from the signatures of the appliances. These features include the durations and time intervals between the various states, existence of repeating parts in a signature, spikes, rapid power changes, ratios between durations of different states. These features complement the generic profile of an appliance, in an attempt to accurately describe an appliance type as a 'bag of features'.

A pattern recognition technique has been developed which, given a profile of an appliance, disaggregates the instances of the appliance in an aggregate signal.

The designed technique is tested on three multi-state appliances, which pose problems to the 'conventional' approaches which only detect steady states. The three targeted appliances are: the dishwasher, the washing machine and the dryer. The system creates the generic profiles and disaggregates instances of all three appliances in aggregate signals with satisfying accuracy in most cases. For two out of the three appliances, the disaggregation was performed on untrained models, with the system managing to successfully generalize.

An important accomplishment is the very small number of detected false positives. The use of a large set of features in order to describe an appliance prevents the detection of false positives in the aggregate signal.

1.4 Outline of the Report

Chapter 2: Presents the background and related work on energy disaggregation, analysis and issues unresolved. Some helpful information of techniques used during this work are also presented.

Chapter 3: Presents the first investigation of the available datasets which are used for this work, as well as the first comparison of features that led to the initial decisions about the approach to be attempted.

Chapter 4: In this section are explained the first tools implemented, as well as the first disaggregation attempt. The prototype designed for the disaggregation of the fridge has influenced the design of the main system. It would be useful (and hopefully interesting!) for the reader to go through this section before continuing with the design of the main system.

Chapter 5: An overview of the developed technique is presented, along with the main functionality of the implemented system and its various modules.

Chapter 6: Describes the process of forming the profile of an appliance. The Profiler module determines the main parts of the appliance's signatures. In this section are also presented the various features observed, their extraction and their incorporation into the profile.

Chapter 7: Describes the pattern recognition technique developed in order to detect patterns in the aggregate signal that match the various aspects of the profile and disaggregate the true operations of the appliance.

Chapter 8: The choice of the classifier is presented and its functionality in the system.

Chapter 9: The system is tested by attempting to disaggregate the three targeted appliances, the dishwasher, the washing machine and the dryer. The results and problems observed are presented and analysed.

Chapter 10: Overall discussion of the results and the achievements after the evaluation. The limitations of the system are also presented. Finally, possible future work for extensions of the technique is proposed.

2 Background and Related Work

2.1 Energy Disaggregation - Why?

When doing the everyday grocery shopping, it is easy for a consumer to compare the prices of the different products and make an informed choice, depending on his needs and budget. While travelling by car, by checking the fuel meters the driver is able to know exactly how much fuel has been consumed since the beginning of his travel, or even how much the instant consumption rises when he accelerates to bypass another car. The case of electricity consumption, however, is not that simple.

In everyday life electricity is taken for granted. It is supplied to homes and businesses and is used by appliances without any indication of the amount consumed. It is in a way invisible. Consumers only get information on their whole-house consumption when they get their electricity bill. That makes it hard to have a feeling about how much electricity is used while, for example, playing a video game. Would the consumer still choose to play it if it was known exactly how much it costs?

Research on the disaggregation of electricity consumption has been performed for many years, with the aim of accurately inferring the consumption of the individual appliances. It attracted, however, much more interest these recent years because of the increasing installation of smart-meters. A smart-meter installed in the main feed of a house measures in regular intervals characteristics of the electricity consumption, for instance the instant real-power consumption, voltage, current etc. Measured characteristics differ between smart-meter models. Many benefits can be gained from a processing technique that can reliably and accurately infer (disaggregate) the individual consumption of each appliance only by processing the whole-house signal.

Figure 2.1: A smart meter installed on the electric main.

Imagine if at the end of the month you would receive an electricity bill with analytical listing of the consumption of each appliance, and you would notice that you spend far more money on lighting than you would have thought. Wouldnt that make you consider the option of changing your bulbs for more efficient ones? Or even try to remember turning off the lights when you leave a room? This kind of changes of consumer behavior has been studied a lot in the literature [9, 10, 11, 12] suggesting that feedback does bear significant results. Depending on the frequency and type of feedback, consumers may optimize their behavior and lower their electricity consumption by up to 20%! In her research [10], Fischer outlines that households would prefer their electricity bills to include clear indication of the various components of the electricity price and she concludes that feedback in order to be most efficient needs to include appliance-specific breakdown. In [12] it is suggested that the most important reason why appliance-by-appliance information leads to greater energy savings is the fact that it enables personalized recommendations. After the identification of the appliances that consume the most, the consumers can be given personalized feedback on how to reduce their consumption.

This kind of information for individual appliances can be acquired by installing a consumption monitor between an appliance and the socket, without the need for any disaggregation techniques. However the cost of these devices is not negligible, especially taking into account the number of appliances in a house. It could add up to significant amounts if a consumer would like to monitor many of them. It also burdens the consumer with the effort for installation. Smart-meters on the other hand can be installed by the utility, without any disruption of the consumer. The recent study [12] also compares other options, like the use of smart-appliances, but concludes that smart-meters in combination with the use of disaggregation algorithms are the most cost-effective way to go.

Disaggregated data can be of great use to other bodies as well. Utilities would be capable of providing better services by suggesting, for instance, the best tariff-plan for each consumer or by informing the consumer of unusual consumption of a particular appliance, which could mean the appliance is malfunctioning. It would also benefit research a great deal, since this information could lead to the development of better standards and more energy efficient appliances.

2.2 Non-Intrusive Load Monitoring

As discussed in the previous section, a method to acquire information on the consumption of an individual appliance is to install a measuring device between the appliance and its socket. This method is called "intrusive monitoring". The second method, the one that will be the focus of the rest of this research, is the disaggregation of data for the total load of a building, acquired by a single measurement device (a smart-meter). Processing that data, it is possible to acquire separate information for the consumption of lighting or the washing machine. This method is far more convenient since it only requires the installation of one meter, which can be installed by the utility, and has lower costs than acquiring a measuring device for each appliance. Generally called **"Non-Intrusive Load Monitoring" (NILM)** or **"Nonintrusive Appliance Load Monitoring" (NALM)**, it is considered the most practical and promising method and has been the focus of research for many years.

2.2.1 Data Collection for NILM

Almost all the smart meters measure the loads real power, current and voltage. Some may measure reactive power, phase of the current and voltage or the power factor. Of major importance is the sampling frequency of the smart meter. Different sampling rates allow for different type of information to be extracted from the aggregate data. Sampling with low frequency (lower than 1Hz) creates the problem of different events might seem like they happen in parallel. For example, if one sample is taken every 5 seconds and a toaster is turned on 3 seconds after the kitchen lamp, the two events will be recorded as if they happened at the same time. This poses one of the major challenges disaggregation techniques have to solve. On the other hand, high frequencies enable the

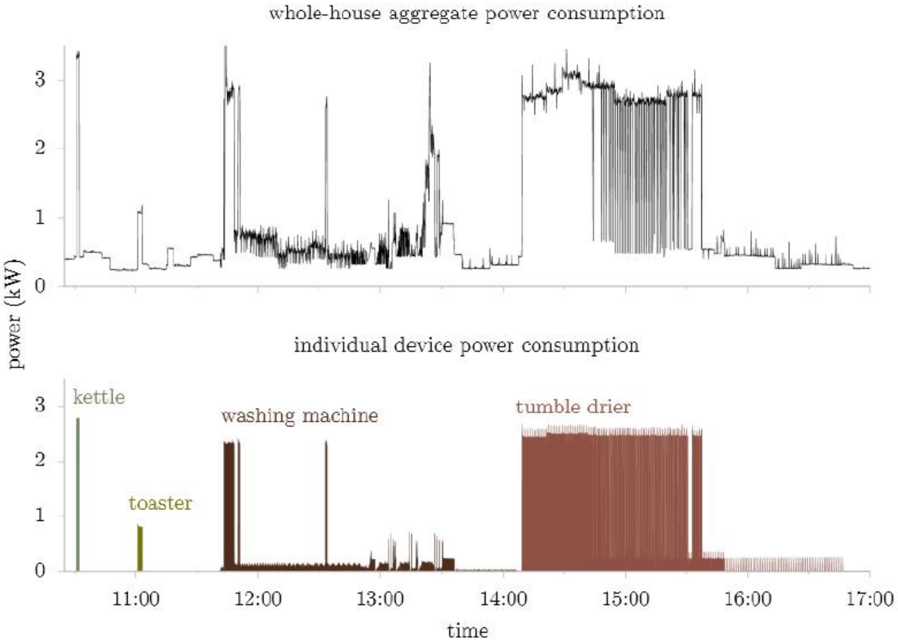


Figure 2.2: Total consumption of a whole house (top) and the individual consumption of four devices (bottom) over the course of an afternoon. Source:[13]

detection of transient events, harmonics and even voltage noise. These techniques will be discussed further below. Unfortunately, technology for sampling at higher frequencies comes with higher cost. Most common smart meters sample up to 1Hz, with the price increasing as the sampling rate goes up. For reference, the smart-meters that are going to be installed during the rollout in the UK are expected to have a sampling frequency of one sample per 5 seconds. Energy meters that sample at a rate of 10 to 100MHz, which are used for voltage noise detection, are usually custom built and expensive [16].

2.2.2 Previous NILM Approaches

The problem of energy disaggregation is not new; researchers have been working on NILM for more than 20 years. There are numerous techniques that have been tried, of which the most important ones will be discussed below. Some very interesting, recent general reviews of the existing techniques can be found at [15, 16].

All of these approaches follow a main principle. They process the data collected by measuring devices and try to extract "**features**" that characterize the devices. A feature can be any distinctive value, behaviour or pattern that can be thought to be distinctive of a device. The disaggregation algorithms then look in the aggregated data for these features, try to recognize them and thus manage to distinguish the separate devices inside the total. The most common feature is probably the transition between steady states in the instantaneous power consumption. For instance, if a device consumes 1kW when it is ON, then if the algorithm finds a power change in the aggregate signal of 1kW, then there is a fair chance that this particular device was turned on at that point. Features will be further discussed in the next sections.

An initial classification of the different disaggregation methods can be done depending on whether they tackle the problem using Pattern Recognition methodologies or they target it as an Optimization problem.

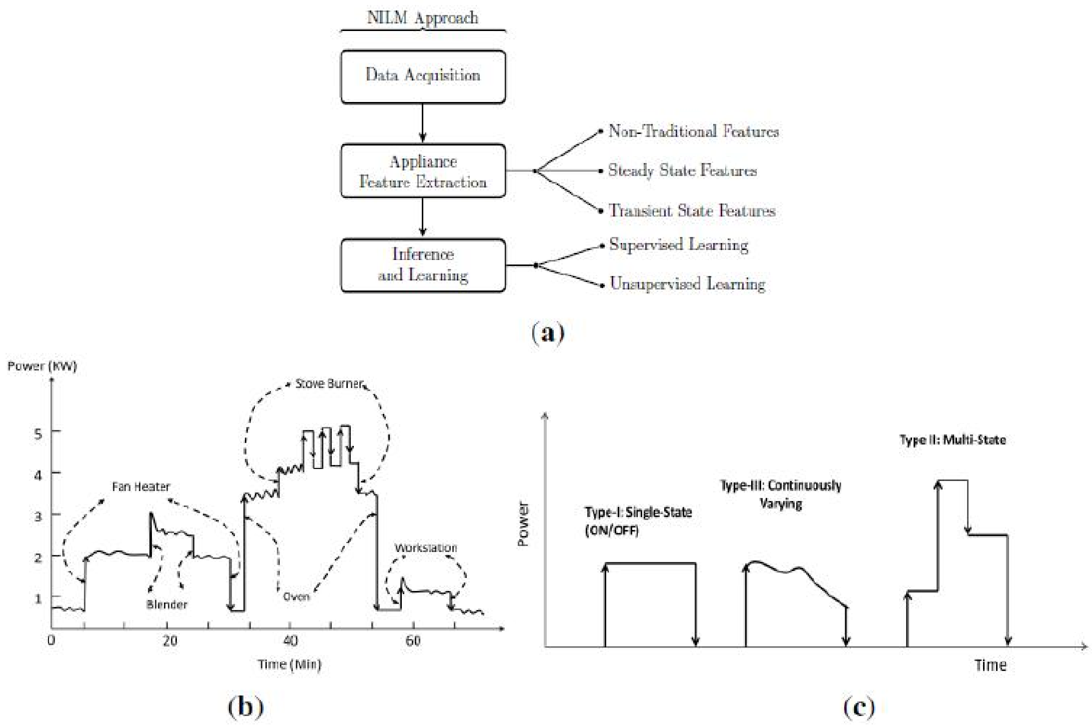


Figure 2.3: a) General framework of a NILM approach. b) The aggregated consumption data obtained from a single point of measurement. c) Different load types based on their energy consumption pattern. Source:[16]

2.2.2.1 Pattern Recognition

An appliance presents some electrical characteristics during its use. Some examples are the value of real and reactive power consumed while it is turned on, a spike at its consumption for a short time after its startup, a slow decay of its power consumption over time, some particular distortion in the voltage or current caused by its internal hardware. All these features can be collected to form a “signature”, a pattern, which is distinctive for a device.

Pattern recognition techniques create a pool of devices and their corresponding signatures and use them to disaggregate the total measured signal. A pattern recognition algorithm tries to detect sets of features in the aggregated signal and then tries to find a pattern within the pool of signatures that matches it close enough. It then proceeds with another set of features, until it has matched all that it can.

Pattern Recognition techniques are the ones that are most frequently used by researchers, mostly because of computational performance in comparison to the alternative, optimization techniques. They can be efficiently used even when the data pool held for known devices is not complete, a major drawback of optimization techniques [16].

Depending on how the techniques create the pool of known devices-signatures, they are further classified as supervised and unsupervised techniques.

Supervised Techniques:

Most attempts on disaggregation have been use supervised methods. In this category, training algorithms are used in order to create the pool of known signatures by processing pre-collected data for each device. For most techniques, the consumption of each type of device has to be recorded prior to the actual disaggregation for a fairly long duration. The training algorithms process that data to identify a distinctive set of features for this device and create the corresponding signature, which is labelled by the researcher. After the pool of signatures has been created, it can be used for the pattern recognition and disaggregation of the total consumption data.

The first and probably the most important work on disaggregation was done by George Hart back in 1984 [17]. For his technique the real power, reactive power and voltage of the load are measured every one second. In order to detect the individual appliances inside the total consumption signal, he used a clustering technique based on the change of power consumed between steady states.

Hart initially identifies three types of appliance models:

- ON/OFF - Two distinct states. The appliance has only one ON state.
- Finite State Machine (FSM) Allows for an arbitrary set of discrete states and state transitions, for example appliances with multiple settings.
- Continuously Variable Infinite number of states, continuous range of operating power levels.

The consumption patterns of these types can be seen in Figure 2.3 (c) as Type I, Type III and Type II respectively.

An edge-detector algorithm passes through the aggregate data and detects all the steady states and the changes between them. The changes of real and reactive power are then clustered together in a two dimensional space. Changes of equal magnitude but opposite sign are matched together, representing the turn on and off cycle of a device. Based on the assumption that different devices consume different amounts of real and reactive power when they are operating at a steady state, the operating cycle of a device can be detected within the aggregate signal by the changes in power consumption. Under another similar assumption, that "distinct states in an FSM appliance model have distinct operating power levels" (Uniqueness Constraint), Hart's algorithm tries to identify FSMs.

Hart's NILM revolutionary technique of course has its limitations. Due to the fact that it uses the changes of power between steady states to identify individual appliances, it cannot handle the "Continuously Variable" type of appliance, mentioned above. Moreover, it cannot distinguish between two different devices if they consume approximately the same power on steady state. The technique is also problematic when more than one devices change states at the same time. His work has later been extended by Laughman et al [23] to incorporate more features than the power consumed, in order to overcome these limitations. This work will be further commented in section 2.2.3 about Transients and Harmonics. Further extension to Hart's work was made by the researchers Cole and Albicki in [34] and [35], where they make use of additional features like edges and slopes in the power signatures to identify particular devices.

An interesting supervised technique seems to be the one developed by Ruzzelli et al [18] using Artificial Neural Networks. A very interesting fact is that they use, among others, the power-factor of an appliance as a feature, allowing their method to distinguish between resistive, capacitive and inductive types. They test their method by trying to disaggregate four types of ON/OFF appliances (kettle, heater, microwave and fridge). They report high accuracy for their method, but from a testing environment where the rest of appliances have a much lower consumption. Moreover, their technique cannot be used for multi-state appliances.

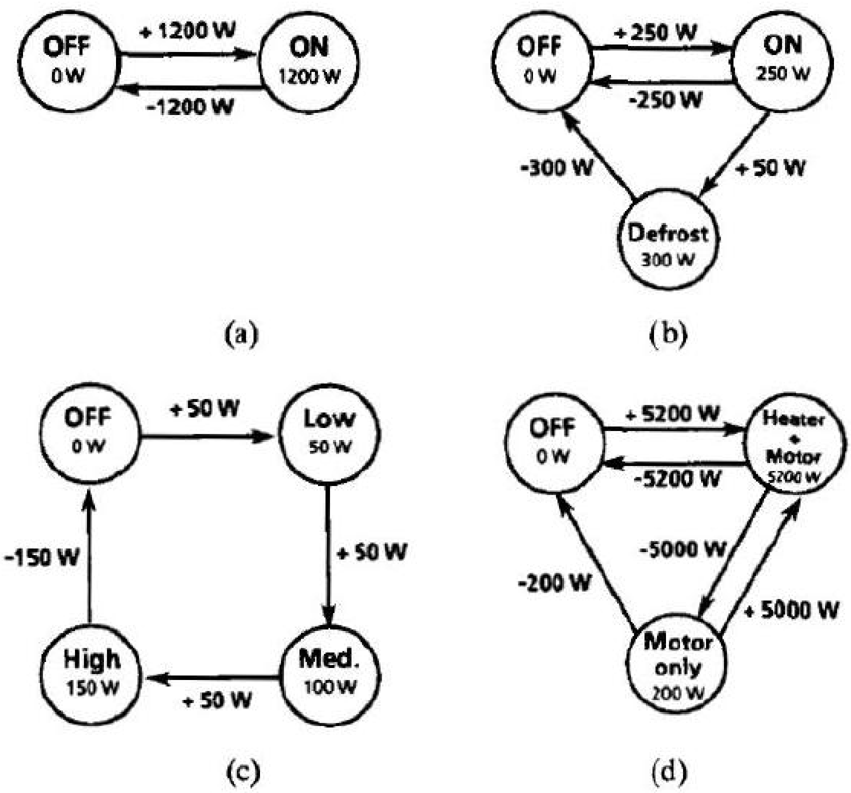


Figure 2.4: Finite State Appliance Models. The circles depict steady states of operation, the edges represent the changes from a steady state to another. a) generic 1200W two state appliance, b) refrigerator with defrost state c) three way lamp, d) clothes dryer. Source:[17]

Finally, a fairly different technique is described in [19], although limited in scope. The technique evaluates certain rules to identify possible time intervals that a water heater or the refrigerator may be operating, in order to compute their power consumption. It also used additional features for the signatures of the devices, for instance the frequency of activation. The sampling rate used was one sample per 16 seconds; however the assumption that two appliances do not activate or deactivate at the same time was made.

Unsupervised Techniques:

Unsupervised learning algorithms, unlike supervised methods, do not require a separate set of measured data for each device in order to process them and learn the devices signatures. Instead, they get trained straight from the aggregate data. This feature of unsupervised learning algorithms has attracted the interest of researchers, in order to avoid the problem of having to collect in advance data for the individual appliances.

A recent and very interesting work following an unsupervised learning approach is presented in [33]. Kim et al used four different Markov models in their technique, which are depicted in Figure 2.5. The FHMM is the initial model, incorporating the states of all the appliances. By the addition of extra features, FHMM is extended to the CFHMM model. FHSMM is another extension of FHMM which alters the probability distributions used for the state occupancy durations of the appliances, in order to make them more accurate. The second and third models are finally merged into the CFHSMM model.

Kim et al tested their method on fairly low frequency data (a sample per 3 seconds), evaluating the three models separately. The forth model comes first, with accuracy approximately 75% for

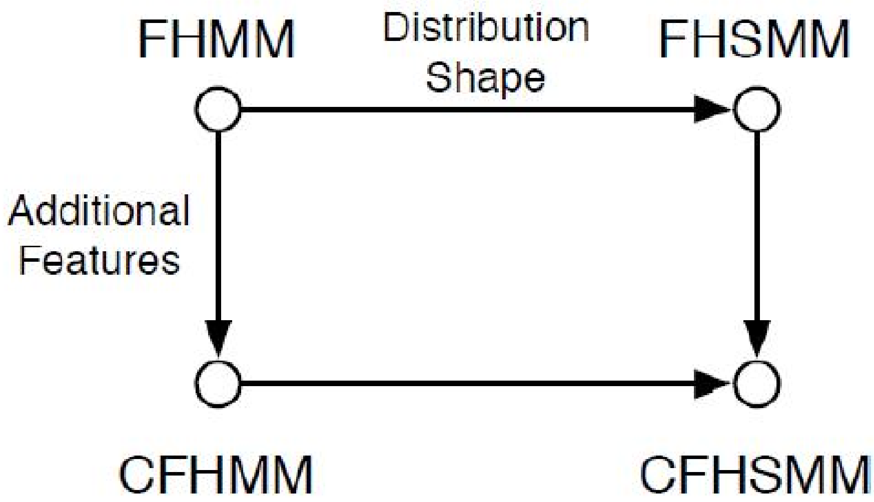


Figure 2.5: Relationships between the various models used by Kim et al. Source:[33]

two houses with eight appliances each. The accuracy of the models decrease as the number of monitored appliances rises. A very important point to notice, though, is that the CFHMM model, which was extended with extra features, performs almost as well as CFHSMM and it does not suffer from accuracy decay as much as the FHMM and FHSMM models, when the number of appliances increases. That finding is an indication of how important is the use of a rich set of features.

The features used by Kim et al are also of great interest. Some of the non-power features used are: time of the day that an appliance is used, duration of use and, finally, a correlation of usage between appliances. For example, there is obviously a dependency between the usage of an Xbox and the TV.

The novel work of Kim et al has its limitations as well. As the number of the appliances increases, the Hidden Markov Models seem to lose in performance rapidly. The maximum number of devices used in the tests was 10. Another drawback of the technique is that it needs to be given the numbers of present appliances before the disaggregation. There is no discussion in the paper on how the algorithm performs if that number changes by the addition of an unknown appliance (eg. if a new radio is bought and installed in the house). Finally, the developed method only works for ON/OFF type of appliances and not with appliances that might have more states of operation.

2.2.2.2 Optimization Optimization techniques, like pattern recognition, make use of known signatures and try to find the optimal combination that produces the aggregate signal. The difference with pattern recognition is that the latter processes and tries to detect in the aggregate signal each device individually. Optimization techniques on the other hand try to find the optimal solution by detecting the occurrences of all the known signatures at the same time.

Hart describes in his work [17] the optimization problem as follows:

$$\hat{a}(t) = \arg\min_{a} |P(t) - \sum_{i=1}^{n} a_i(t) P_i$$

where $P(t)$ is the total consumption, P_i a p-vector of the power that the $i-th$ device consumes when it is operating, describes the state of the $i-th$ device, $\hat{a}(t)$ is a vector with all the states of the i devices.

Solving the optimization problem becomes particularly computational expensive as the number of appliances increases [16]. Hart suggests it is "computationally intractable". Approaches that have been tried in order to reduce the complexity of the problem include genetic algorithms and integer programming. Moreover, optimization techniques suffer from one more drawback in comparison to pattern recognition techniques. The fact that all the devices cannot always be known complicates the problem further, because the optimization algorithm tries to find the optimal way to combine all the known signatures to reproduce the aggregate signal [16]. That factor does not affect pattern matching techniques since they try to detect the known signatures in the composite load one by one at a time.

M. Baranski and J. Voss in [20] developed an interesting method that uses a genetic algorithm to combine clusters to create finite state machine models. Some very interesting research was done in [21] and [22]. The authors combine the use of different algorithms, both optimization (integer programming, genetic algorithms) and pattern matching (Artificial Neural Networks). A "Committee Decision Mechanism", as they name it, processes the output of the previous algorithms and computes the optimal final output. It achieves that by using either the Most Common Occurrence, Least Unified Residue or the Maximum-Likelihood Estimator. Their simulation tests show that their use of CDM increases the accuracy over the accuracy of every individual algorithm by approximately 10%. The reported overall accuracy is impressive, reaching 90%. However, it seems as if their overall accuracy metric does not take into account false positives. Another distinctive point of their work is that they use a very rich pool of features which among others includes harmonics and transients.

2.2.3 Disaggregation at high sampling rates - Transients, Harmonics, Noise

In some cases, different kind of devices might have identical power consumption when they are on. That led many researchers to use "microscopic" features, like short transients and harmonics in order to achieve higher accuracy of disaggregation.

Hart discussed the use of "Transient Signatures" in [17], although they were not used in his original disaggregation work as they are harder to detect. Transient waveforms differ in shape, duration and size; hence they can be used as features of a devices signature. Hart identified three types of transients, two of which are characteristics of motors. The third one is found in several appliances, varies a lot and is shorter than one or two voltage cycles. In [14] Zeifman and Roth suggest that a sampling rate of at least 1kHz is required to detect the harmonics and have distinctive transient signatures.

The work in [21, 22], also commented in the previous section, manages to make use of microscopic features, harmonics and transients, using a sampling rate of at least 12kHz. Laughman et al in [23] use harmonic analysis during transient events, with a sampling frequency of 8kHz. Making use of Fourier transformation, they compute the "spectral envelops" that summarize the harmonic contents of the system. The accuracy is not exactly known and in order to perform well it needs excessive training beforehand. Norford and Leeb in [24] and Chang et al in [26] also made use of startup transients in a similar technique.

A novel and interesting approach that makes use of high-frequncy data was presented by Patel et al in [25]. This method monitors the electrical noise generated by the appliances on startup events and steady state operation due to their electronic components. It required the use of a custom device that is plugged in a socket, excessive training and sampling frequency of 500kHz.

The above techniques are generally considered accurate and many researchers agree that the use of microscopic features should complement the techniques of the previous section to improve accuracy. As the most commonly used and cheap smart-meters do not have the capability to measure at that high frequencies, however, it is preferable to avoid the use of these techniques if possible, at least until measuring devices sampling at higher frequencies become cheaper and more widely used.

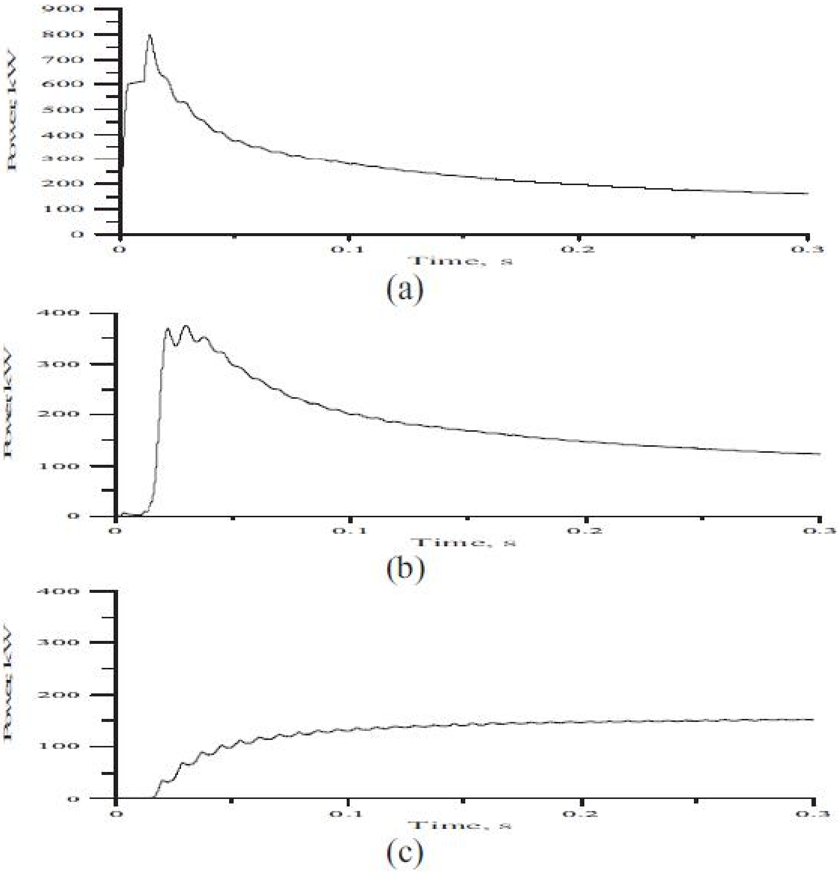


Figure 2.6: Turn-on real-power transients for three different types of motors. Source:[26]

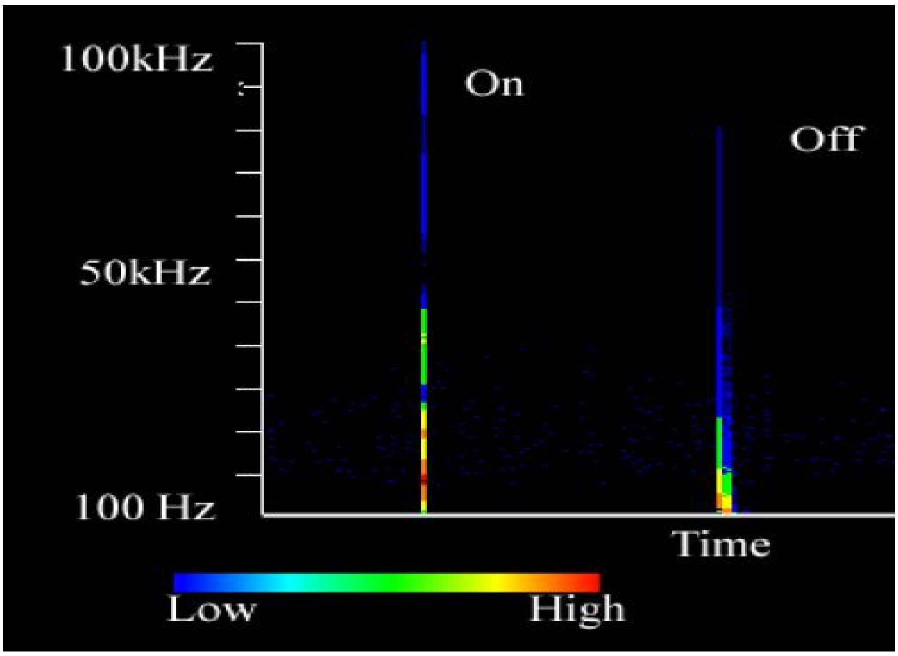


Figure 2.7: Frequency spectrum of a particular light switch being toggled (on and off events). The on and off events are different enough to be distinguished. Source:[25]

2.2.4 Issues

Even though research on disaggregation has lasted for over 20 years, still there is no developed algorithm that has reported high accuracy and could be widely adopted in the domestic sector. Some techniques were developed and tested for only a few types of appliances, some require special equipment for sampling at very high frequencies, others have not reported accuracy. In this section are discussed some of the major issues that hold research back.

Lack of publicly available datasets:

The development of disaggregation techniques is all about processing of data. From determining the features to be used and training of the algorithm all the way to testing the results against the ground truth, a rich data is essential. Energy measurements take time to collect and can block a research if they are unavailable. But this is not the only problem. Typically, the developed algorithms were tested on the same data that they were trained from; same devices but in different operation cycle. However, devices of the same type but of different manufacturer might present different electrical characteristics. So, there is no assurance that the developed algorithms would be as accurate disaggregating the consumption of other households.

It has been only recently that the first public databases became available, with the first initiative coming from the MIT with the REDD dataset [27]. More datasets became publicly available afterwards, like the ones in [28, 29, 30]. Still a lot of work can be done towards a dataset complete enough that can serve as a reference point. It would allow developed algorithms to be tested against a common target and allow for proper comparison of their efficiency.

Generic Signatures:

From the literature it became obvious that there is no certain set of features that can be used to disaggregate all the different types of appliances yet. Harts work [17] was a great start, but the use of power alone to distinguish between appliances has the drawback that it is not possible to differentiate two different appliances that consume same power levels. Later efforts introduced the use of several other features. Attempts like those presented in [19, 23, 33, 34, 35] prove how the use of extra features can improve accuracy of an algorithm. Still, it seems there are not many features determined that are specific for certain appliances so that they could be used to disaggregate them with high level of accuracy.
In addition, as mentioned above, since in most research efforts the algorithms were trained and tested using the same dataset, it could be the case that the algorithm would not perform well disaggregating a new, unknown aggregate signal. This is based on the fact that different models of the same appliance type may present different electrical behaviour, for instance power consumption or duration of operation. The design of generic signatures through the correct set of features, which would be detected independently of testing conditions and appliance model or size, would be a great step forward.

Standardized performance metrics:

Finally, even though research on disaggregation has been going on for years, there are no established performance metrics. That makes it really hard to compare the developed techniques and distinguish between efficient and not efficient approaches.

Some algorithms are developed targeting certain types of devices and their accuracy is evaluated on a per device basis. Some others use overall accuracy metrics dependent on the number of device instances successfully detected, or the percent of the total power consumption that was successfully disaggregated. However how does one treat false positives in the accuracy evaluation?

In their very interesting work [22], J. Liang et al have used three different metrics for the evaluation of accuracy. These metrics are: Detection Accuracy, Disaggregation Accuracy, Overall Accuracy. This helps evaluate independently the detection module and the disaggregation module of their technique. However, their overall accuracy metric seems not to take into account false positives. Also, it is difficult to infer relations between the different metrics. Another example, in [26] the results from four different case studies are reported with significant differences in accuracy. However it is not possible to determine if the system is overall efficient.

2.3 Optical Character Recognition

This work has been initially inspired by techniques used in Optical Character Recognition. Without getting into too much detail, an idea of the techniques used in OCR might provide an initial intuition into the approach.

Optical Character Recognition techniques analyse a handwritten, typewritten or printed character in order to classify it correctly - to make the computer understand which character of the alphabet or number it is. Most modern approaches process the given character and attempt to extract several features from it. Many different techniques are being used, but all of them aim at analysing the shape of the given character. When a set of features has been extracted, it is compared with a stored database of feature-vectors which represent the known characters. The already known and stored feature-vectors constitute the profiles of the alphabet characters and numbers. If a close match is found between the features extracted from the given character and one of the profiled ones, the scanned character is classified.

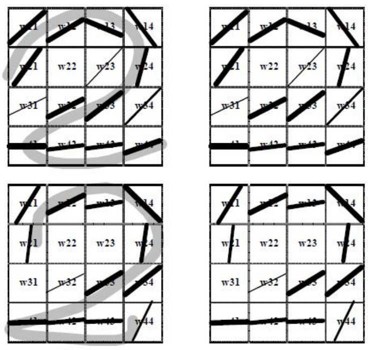
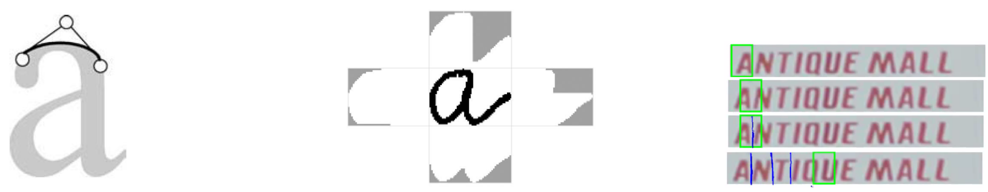


Figure 2.8: Example of parametrization of two characters. Source: [38]

Several techniques are used in order to extract features that can effectively characterize a character, frequently used in combination with each other in order to achieve higher accuracy. In order to define the shape of a given character, some make use of histograms, analysing the distribution of the points that make up a character from different projections. Others aim at decomposing the character into smaller parts, for instance lines, edges, loops, curves and determining their relative position. Such examples are shown in Figure 2.9. Additional features extracted are the direction and orientation, as well as topological features, such as start, end and cross points [37].



(a) (b) (c)

Figure 2.9: (a) Detection of curves of a character with the Bezier method. Source: [37] (b) Histogram of different projections of character 'a'. Source: [36] (c) A sliding window scans an image in order to detect the areas that contain a character.

Finally, a technique used especially in Photo OCR which inspired a part of the disaggregation technique is the 'Sliding Window Technique'. A window scans the image in order to detect parts that contain characters (see Figure 2.9c). After these parts are detected, the main processing and classification of the character takes place.

2.4 Gaussian Mixture Model

A Gaussian Mixture Model (GMM) is a probability density function which can be interpreted as being derived by a weighted sum of a finite number of Gaussian density functions. Gaussian mixtures are capable of representing a large variety of sample distributions by forming smooth approximations. Apart from the ability to fit arbitrarily shaped distributions, GMMs are also frequently used for modelling underlying hidden classes in the data by their different component density functions [39].

Given M component Gaussian density functions $g_1(x), ..., g_M(x)$, with corresponding means and standard deviations $(\mu_1, \sigma_1), ..., (\mu_M, \sigma_M)$, the density function of the Gaussian Mixture Model is given by the equation:

$$p(x|\lambda) = \sum_{i=1}^{M} w_i g(x|\mu_i, \sigma_i),$$

where λ represents the set of parameters $\lambda = w_i, \mu_i, \sigma_i, i = 1, ..., M$ and the components' weights $w_i, ..., w_M$ satisfy the constraint:

$$\sum_{i=1}^{M} w_i = 1$$

The density function of each Gaussian component function is of the form:

$$g_i(x) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{\frac{-(x-\mu_i)^2}{2\sigma_i^2}}$$

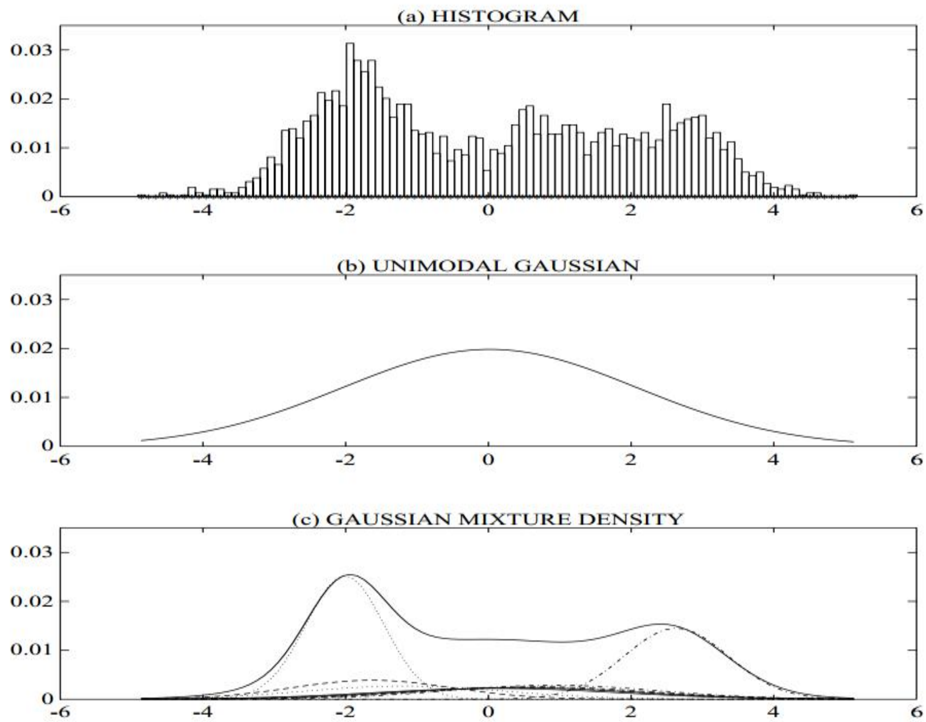


Figure 2.10: Comparison of distribution modelling (a) histogram of samples, (b) modelling the data with a simple Gaussian function, (c) fitting a Gaussian Mixture Model. Source: [39]

Given a set of data, in order to form the Gaussian Mixture that best matches the data it is needed to compute the set of parameters λ. The most common algorithm for the solution of that problem is the Expectation-Maximization algorithm. It is an iterative optimization method that, beginning with an initial λ, it computes for each point in the data the probability of being generated by each component of the model. The set of parameters λ is tweaked in order to maximize the likelihood of the points. The new λ becomes the initial λ of the next iteration and this process continues until convergence is achieved. Estimation-Maximization is a well studied algorithm for which information can be found from various sources, such as [41], for that reason it will not be analysed here in further detail.

2.5 Support Vector Machines

The Support-Vector-Machine is a state of the art supervised binary classification technique, being widely used for its high accuracy, even when dealing with high-dimensional data. SVMs have been introduced in 1992 in [43]. It is essentially a mathematical algorithm for detecting the optimal hyperplane that separates data in a multi dimensional feature space.

In order for an SVM to perform classification, it essentially separates the area of the feature space into two: the area of positive and the area of negative classification. The plane that separates the two areas is the hyperplane. In a two-dimensional feature space, such separation can be performed simply by a line. In a three dimensional feature space, a plane is needed for the separation. The hyperplane is the analogous in a multi-feature space. This hyperplane is also referred to as decision boundary.

One of the main attributes that makes SVMs so efficient and popular is that fact that it can separate non-linear separable data using linear methods. The way it is performed is through the kernel function. The kernel function, by convolution of the dot-product, is able to add dimensions to the data. It is thus possible to project non linearly separable data to higher dimensional space, where they are linearly separable. This trick, also known as 'the kernel trick', can be understood by referring to figure 2.11. Different kernel functions project the data in different multidimensional-spaces, with the efficiency depending in a smart choice of a kernel function. Further analysis of the SVMs can be found in [43] and [44], but it is not required from the reader in order to follow the rest of this work.

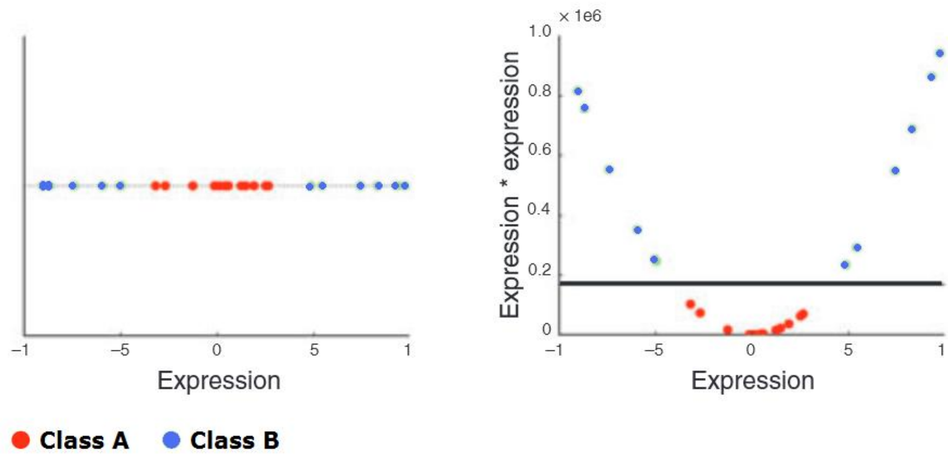


Figure 2.11: An SVM is able to separate non linearly separable data using linear methods, through the use of the 'kernel trick'. For instance, the data in the left picture cannot be separated in the two-dimensional feature space linearly (by a single line). With the use of a second degree kernel function, the data are projected to a four-dimensional space, in which they can now be linearly separated. If the data are now projected back to the two-dimensional space, the linear boundary will have the form of a curve.

3 Exploring the Data

One of the main problems that research attempts on energy disaggregation stumble upon is the little availability of data. That may result in various inefficiencies, like the design of a technique which cannot generalize well enough on different models of an appliance, if for example the design is based only on one model. From this starting point, before an attempt even begins, the trustworthiness of the final result and its evaluation may be lowered if the data are limited and the designed system is bound to be trained and tested on the same data.

To avoid such complications, the first step in this work was to collect the needed datasets, check for inefficiencies, consider for what types of appliances the available data are enough and finally decide what appliances to target.

3.1 Available Datasets

The first step was to get access to any available data. A rich dataset has been collected by co-supervisor of this project, Jack Kelly, as part of his PhD research. It is also very fortunate that fairly recently some datasets became publicly available. The ones used in this work are described in [27, 28, 29, 30] and some of their characteristics are presented in Table 1. These datasets were the ones found with a sampling rate within or close to the range of interest, 0.1-1 Hz, and were the ones processed in order to find patterns and features of interest, training of the algorithms and, finally, evaluation of the results.

Table 1: Available Datasets

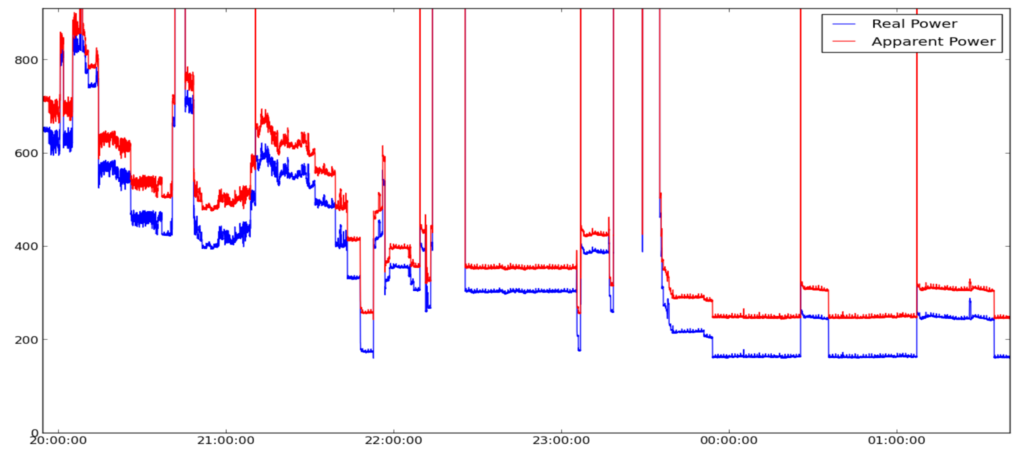
Dataset	Signals	Sampling Period	Apparent Power (S)	Real Power (P)
Jack Kelly‘s	Aggregate	1s/6s	Y/Y	Y/-
	Appliances	6s	-	Y
REDD	Aggregate	1s	Y	-
	Appliances	1s	Y	-
Tracebase	Aggregate	-	-	-
	Appliances	2s	-	Y
Smart*	Aggregate	1s	Y	Y
	Appliances	15-25s	-	Y
AMPds	Aggregate	1m	Y	Y
	Appliances	1m	Y	Y

Sampling frequency

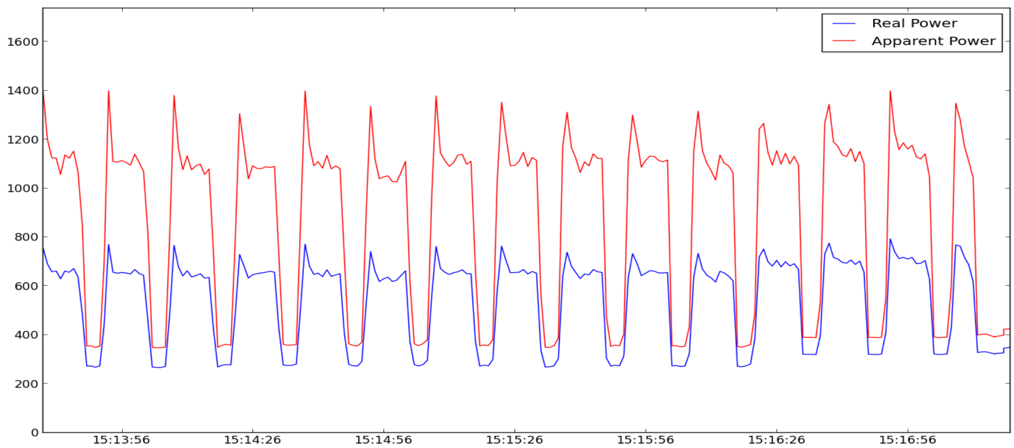
This work aims to utilize data from smart meters sampled within the frequency range of 0.1Hz to 1Hz. The data from Smart* and AMPd datasets are measured at lower sampling rates, not allowing the detection of features like spikes or rapid power changes. They are still valuable, however, for comparing the signatures of different models in order to spot patterns based on states with longer durations.

Active and apparent power

An important issue is that different types of smart-meters do not all measure the same type of values. For instance, the REDD database is created by recording apparent power (S). The Tracebase dataset is of real power (P) measurements. The sensors used by co-supervisor of this project, Jack Kelly, measure both apparent and real power for the aggregate signal (with different sampling rates), but for most of the appliances it is real power that is being monitored.



(a)



(b)

Figure 3.1: Difference between real power (blue) and apparent power (red) is usually in the range of 30-100W (a). The greatest difference was observed during the washing machine operation (b). The magnitude of the rapid changes created by the motor's operation may have over double the size when measured in apparent power than in real power.

An inspection of the aggregate signals available for both real and apparent power revealed two important points. Apparent power is constantly greater than real power by a varying amount. In most stable states, this amount is between 30-100W (Figure 3.1(a)). If small appliances are to be detected, this would make a difference. Most of this work however was focused on large appliances where this change does not complicate much their detection. It does however make a difference when calculating the overall power consumption of the appliance, with the consumption of apparent power being greater. The greatest difference between real and apparent power was noticed during the operation of the washing machine's motor (Figure 3.1(b)). The inductive nature of the motor makes the magnitude of the rapid changes being much greater in apparent power than in real power, fact that needs to be taken into account when calculating the total power consumed. Overall, features like state transitions and spikes seem to be detectable in both cases. Worth noting is that the top power value reached by a spike transition seems to be quite greater when observed in apparent power measurements (even greater than 30%). Although this magnitude was not used throughout this project, it could be a point worth of attention for future work.

3.2 Appliances of Interest

The next step was to investigate the common domestic appliances, check the availability of data for each appliance type and form an initial decision on which to target. The most energy consuming domestic appliances are presented in Figure 3.2. Since the ultimate goal of energy disaggregation attempts is to classify correctly as large percentage of the total household consumption as possible, these appliances were the starting point.

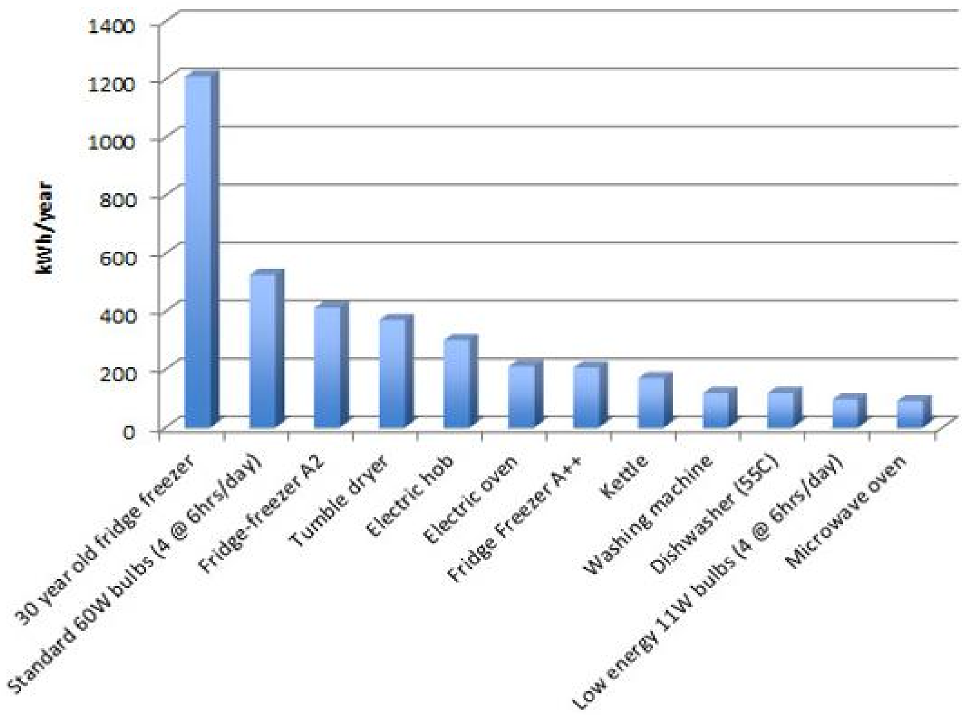


Figure 3.2: Energy use by various appliances. Source: [47]

3.2.1 First Investigation for Features

Most of these appliances operate at fairly similar power levels, as shown in Figure 3.7a. This fact reveals early on one of the difficulties faced: the different models of an appliance type operate in power ranges that overlap with models of appliances of another type. A disaggregation technique that would only use the magnitude of power changes to detect an appliance would be confused between those appliances with similar power consumption. Other characteristics are needed to be found in order to detect one accurately, separating it from the rest.

Figure 3.7b shows the durations of appliances' operations that have been observed in the datasets available. A comparison of the distribution reveals a first clue: the usual operation of these appliances can be used to separate between them. For instance, the microwave is usually for a very short duration, whereas the washing machine usually operates the longest. Further inspection revealed more interesting patterns in the operation of some of those appliances. These patterns for the dishwasher, the washing machine and the dryer can be observed in the figures 3.4, 3.5 and 3.6. The available datasets include a fair amount of different models for these three appliances (twelve models of dishwashers, eleven of washing machines, seven models of dryers) and thus they were the ones that were mainly used in the next steps of this work for development and evaluation.

Top Loader washing machine

Another type of washing machine than the one described above was observed in the REDD and Smart* datasets. The power signature of this type can be seen in Figure 3.3 and differs significantly from the signatures found in the rest of the data. This is due to the different operation of the top-loader washing machines used mainly in the USA, which have a shorter operation cycle and usually function with water that is heated before the insertion to the washing machine, by an external water-heater. This type of washing machine has not been focused upon during this work.

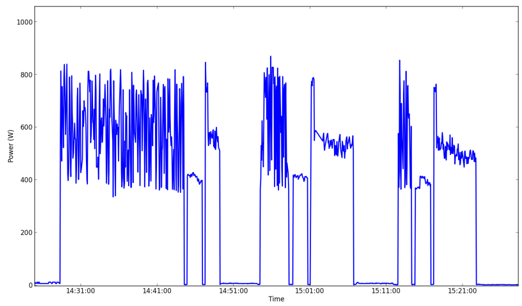


Figure 3.3: Signatures of washing machines found in the REDD dataset.

3.2.2 First intuition into the approach

The initial findings presented above suggest that the signatures of some appliances can be modelled as 'bags of features'. The detection of enough of those features would indicate the identity of the appliance, regardless of the model and size (smaller and bigger washing machines all present rapid power changes for instance). On the other hand, the absence of some of them would indicate that this is not the correct appliance and thus avoid the false detection.

It is also interesting to notice that these appliances targeted are multi-state appliances. This type of appliance is hard or impossible to detect with some of the previous attempts reviewed in the Related Work section, due to the complexity of the signatures. If however these signatures are broken into smaller parts, features are extracted and the signatures are treated as a set of those, it is exactly this complexity that enables their detection. This approach is presented further in the next sections.

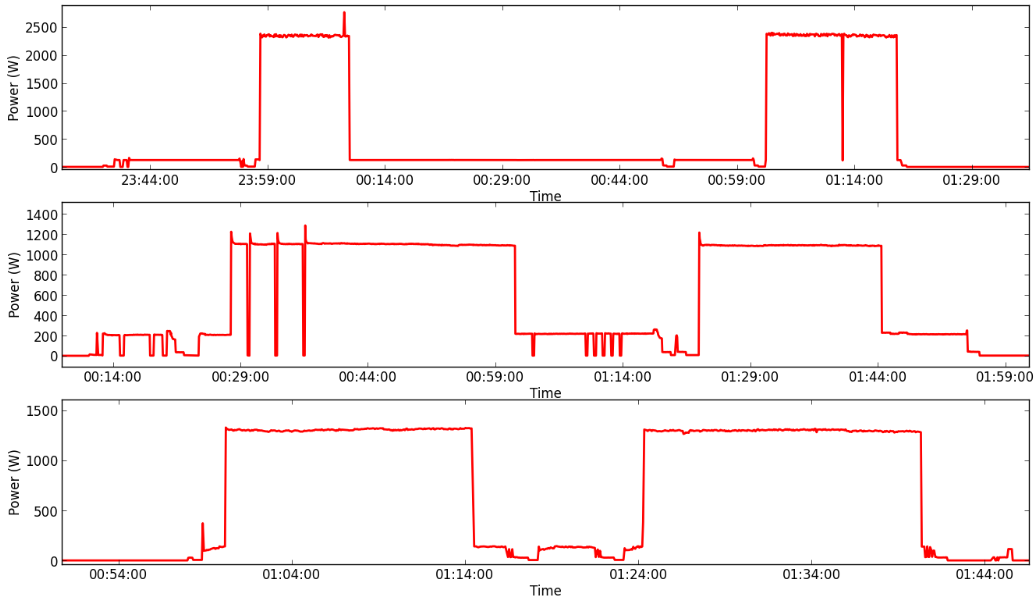


Figure 3.4: Signatures of three different models of dishwashers. During a dishwasher's operation the heating element turns on usually twice. The heating element is the one that consumes high amount of power. The two heating cycles to warm up the water usually have similar durations, hence the similar widths of the two 'blocks' in the signatures. During the rest of the cycle other mechanisms of the dishwasher, like the circulating motor, turn on and off and consume a smaller amount of power.

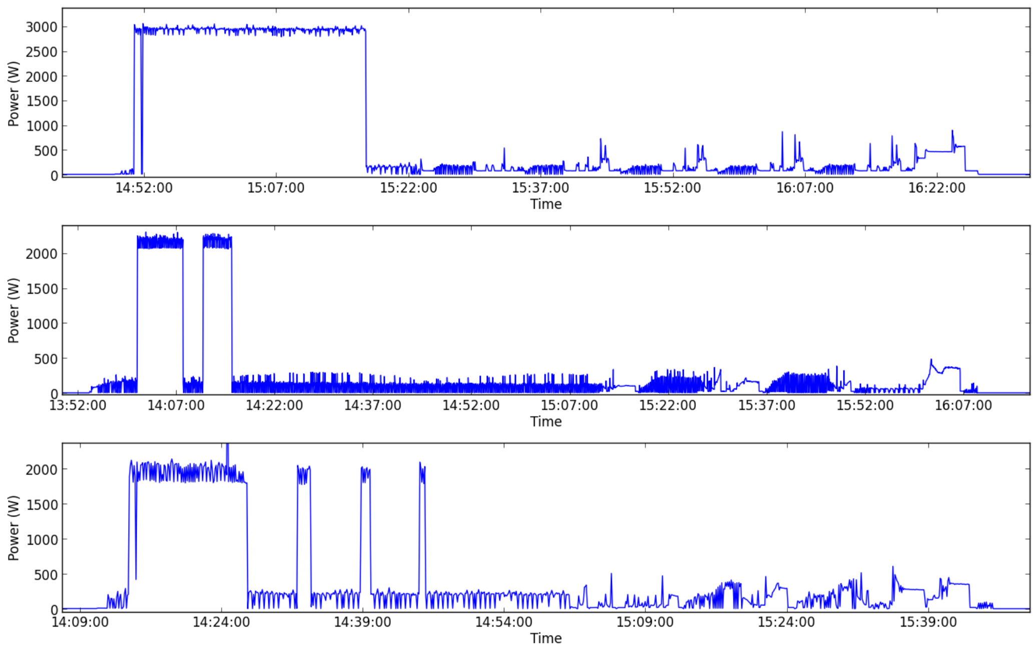


Figure 3.5: Signatures of three different models of washing machines. The washing machine's heating element turns on near the beginning of the operation and consumes high amount of power. In some cases it turns on and off again a few times if the water's temperature decreases. After the main wash, rinsing cycles occur, with a final fast spin in the end of the operation. The washing machine's motor creates rapid power changes throughout the operation.

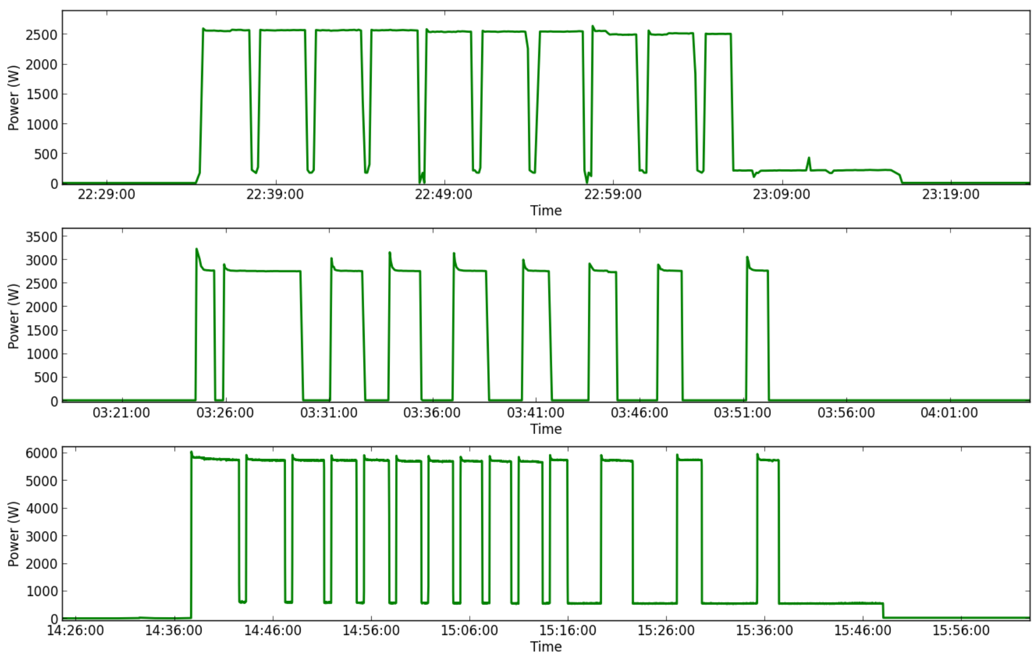
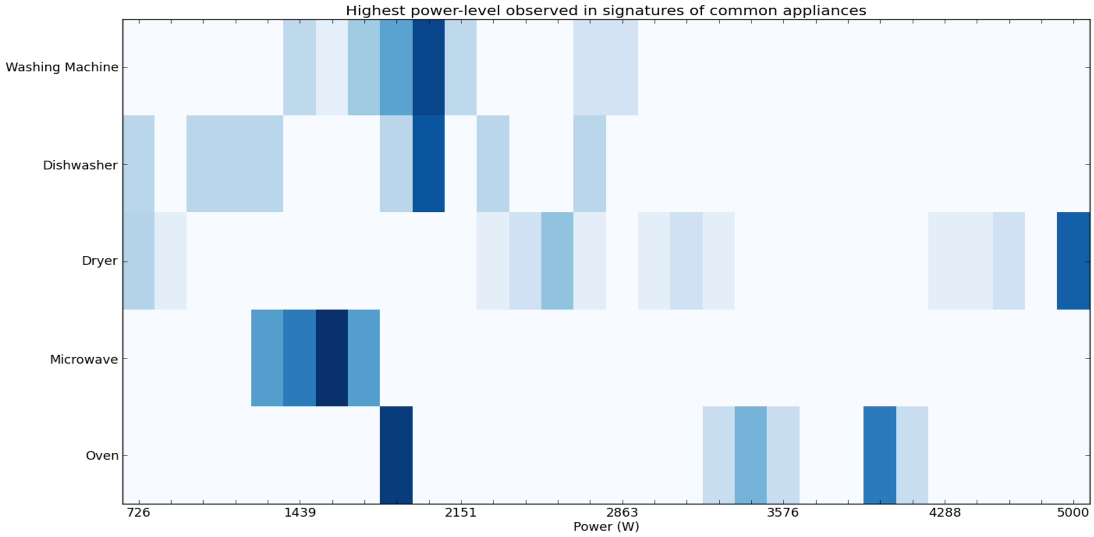


Figure 3.6: Signatures of three different models of dryers. The dryer's operation is mostly a heating element turning on and off several times in frequent time intervals. The heating element usually presents a spike in its power consumption when it first turns on, which can be observed at sampling rates close to 1 second.



(a)

(b)

Figure 3.7: Highest power steady state (a) and durations of one operation (b) for common appliances, as observed in the available datasets. The rows correspond from top to bottom: washing machine, dishwasher, dryer, microwave, oven.

4 Implementing the First Tools - First Disaggregation Attempt

Before any further attempt to process the data and the appliance signatures would be possible, it was necessary to implement the first tools that would detect the steady states, the power changes, the spikes and the rapid power changes in the data. This step was also necessary in order to get acquainted with the different datasets and spot early on any inefficiencies of the measured data.

The second part of this section describes the first attempt in disaggregating an appliance, the fridge. Although the fridge is rather different than the multi-state appliances focused for the rest of this research, this attempt helped form part of the final technique. For that reason it is presented here, in order to give the reader a first idea of the disaggregation process.

4.1 First layer feature detectors

Whenever an appliance turns on or off a change can be observed in the measured power signal. For multi-state appliances, power changes are observed even during an appliance's operation whenever one of its mechanical parts changes state, for instance a heating element or a motor. For the time periods that an appliance does not change state, the power signal is fairly steady.

The First Layer Feature Detectors module that has been implemented consists of the tools that are used in order to detect and process these power changes and steady states in the data.

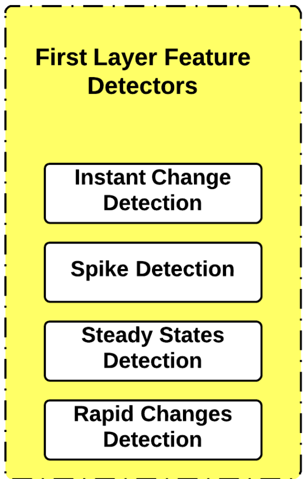


Figure 4.1: First layer feature detectors module.

4.1.1 Steady state detection

Even for periods of time that an appliance does not change state, the power signal is not as stable as one may think. Small changes in the voltage supply create changes in the power signal. Small changes in the state of the electronic parts of an appliance may also create small changes in the power signal. In order to detect the main steady states during an appliance's operation, a threshold proportional to the average power of the state has been used. Any change smaller than that threshold does not 'break' a steady state (Figure 4.2).

4.1.2 Normal and Spike power changes detection

A transition between two operational states of an appliance creates a power change, for instance when an appliance that consumes 1000W turns on a positive power change of 1000W is observed. A negative power change of -1000W is observed when the appliance turns off. In the case of multi-state machines, these kind of changes are observed during its operation whenever one of their individual mechanical or electronic parts change state. Some of those changes are instant whereas some have the form of a spike. Spikes are created mostly by resistive elements (eg heating elements) when they turn on, caused by the fact that when the element is initially cold, its electrical resistance is lower and so it draws greater amount of power. As its temperature increases a few seconds afterwards, the electrical resistance of the element increases, its power consumption decreases and reaches a level where it then remains stable.

The detection of normal power changes is achieved by detecting continuous increase (positive changes) or decrease (negative change) in the signal. Although by naked eye it does seem like those changes are defined by just two points, it is not always the case. Depending on the sampling

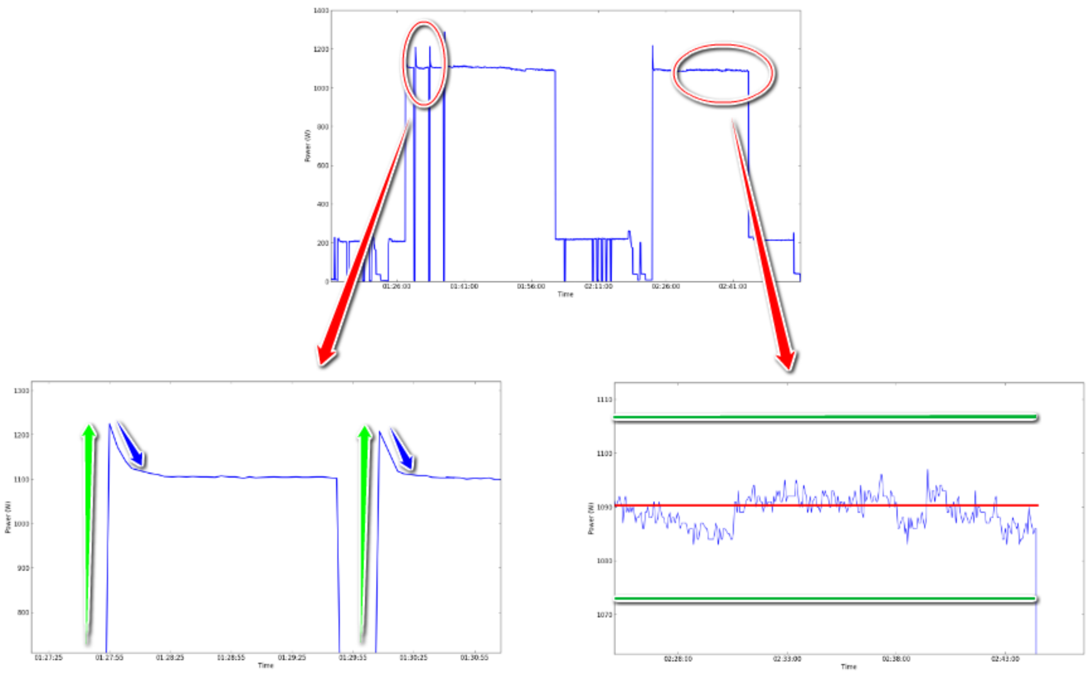


Figure 4.2: Top: Signature of a dishwasher , Bottom Left: ascending and descending phases of spikes, Bottom Right: detected steady state (red) and the threshold of power change (green).

frequency, they might last for more than 1 sampling period. Detecting the spikes was accomplished in a similar fashion, with the detection of continuous increase in the power (ascending phase), followed by a decrease (descending phase). See Figure 4.2.

4.1.3 Rapid changes detection

The operation of some appliances creates rapid power changes in the electricity signal. For instance in the case of a washing machine, such changes are caused by the operation of its rotating motor. Periods of time that such changes occur can be a distinctive characteristic of these appliances' operation. Their detection is also of importance because these rapid changes add 'noise' in the detection of other features. If for instance these changes have a magnitude of 100W, they could be mistakenly detected as the power change created by a fridge that turns on.

The detection and elimination of these changes was based on their fairly high frequency. The detector goes through the data and looks for periods where oscillations occur with a short enough cycle (as if looking for a high frequency sinusoidal signal). After the detection of such periods, the changes are removed from the data, smoothing it for further processing (Figure 4.3).

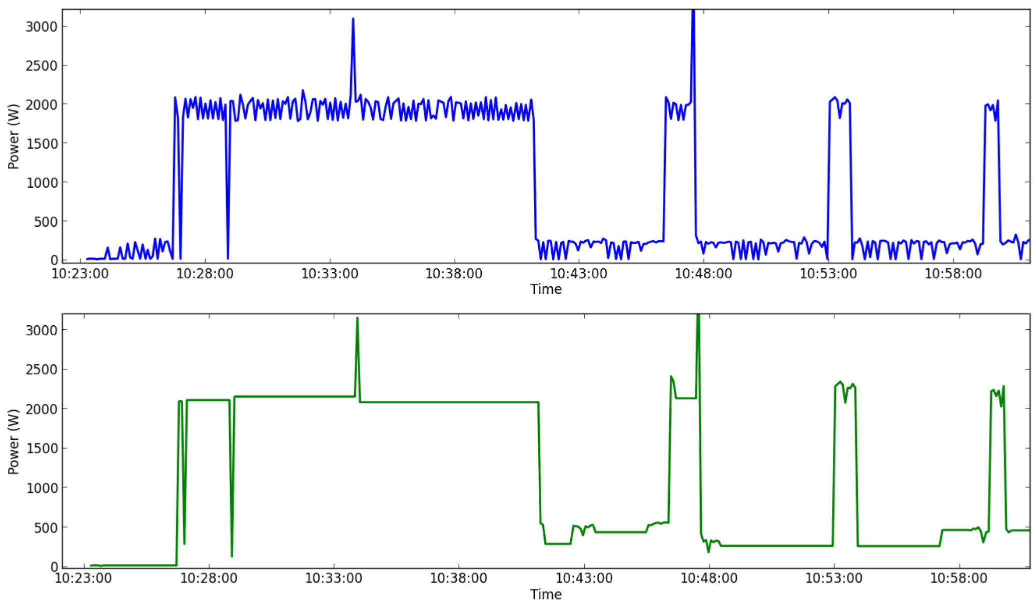


Figure 4.3: Period of rapid power changes occur during the operation of a washing machine. After their detection the 'noise' is removed.

4.2 First disaggregation attempt

In order to get acquainted with the aggregate signal and identify early the implications of the disaggregation process, it was decided to gain some initial hands-on experience by attempting to detect the operation of a fairly simple appliance in the aggregate. The design of the prototype as well as the the problems encountered during its implementation helped form the final system and spot inefficiencies of the approach that should be avoided in the next stages, where multi-state appliances are focused.

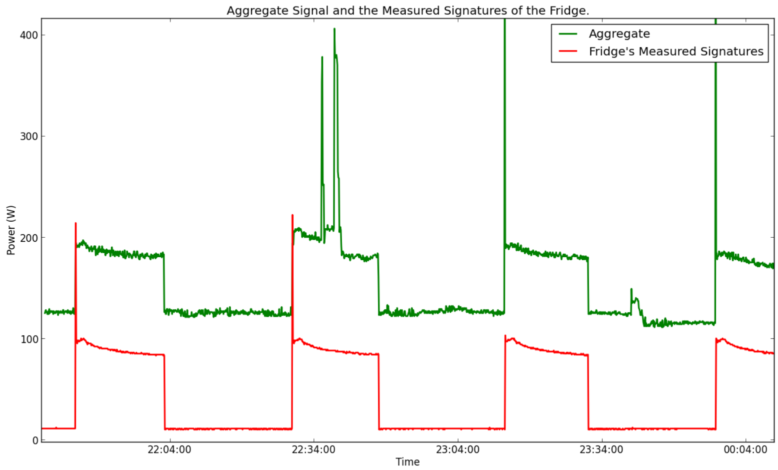


Figure 4.4: Aggregate signal (green) and four signatures of the fridge (red).

4.2.1 The Fridge

Although the refrigerator has been the target of various previous and rather successful attempts, for instance the work in [19], it has a characteristic that made it an interesting choice for this experimental prototype.

The fridge is a one-state appliance, turning ON and OFF in regular intervals throughout the day in order to keep its compartments in the desired temperature (Figure 4.4). As shown in Figure 4.9c, the power consumed during its operation is fairly low, between 80W and 200W. Power changes of such magnitude can be caused by various other appliances, making its detection problematic if based only on that feature. The additional characteristic of frequent operations however (see Figure 4.9b) makes it an exceptional choice for the first experimental design of a technique that detects appliances as a 'bag of features'.

4.2.2 Design and Implementation

The prototype system designed, presented in Figure 4.5, consists of two main modules. The Profiler module takes as input the measured data with the signatures of the various models of fridges. It processes all the signatures and gets 'trained' on the power that the various fridges consume and the duration of their operation. This process is done only once. The output is the 'profile' of the fridge appliance, which is provided to the Disaggregation module in order for it to find relevant power changes in the aggregate signal.

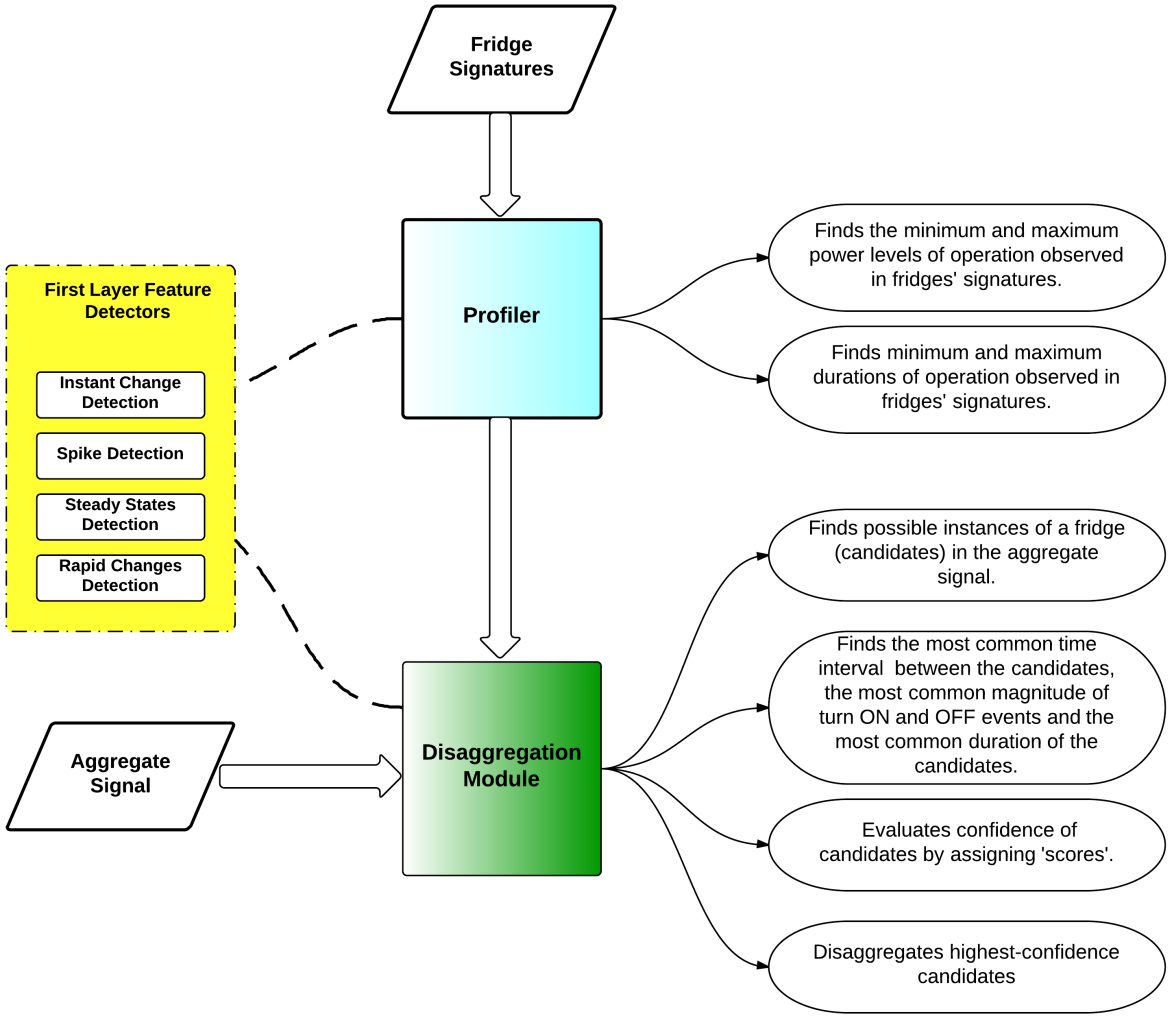


Figure 4.5: Design of the prototype system.

The Disaggregation module finds pairs of positive and negative changes in the aggregate with magnitude and time distance within the limits provided by the input (the profile). These pairs are potential candidates for being the turn ON and OFF events of a fridge's operation.

Afterwards, it is attempted to infer the characteristics of the particular model of fridge, the one present in the aggregate signal. This is achieved by comparing the various candidates and finding the most common magnitude of positive and negative power changes, the most common duration of operation, as well as the most common time interval between two consecutive operations. As it is explained further below, this step is based on the fact that the fridge operates several times throughout a day and these characteristics are fairly stable for several operation cycles (See figure 4.9 at the end of this section.). Finally, the candidates with characteristics closest to the ones found are disaggregated as being the true fridge operation cycles. This process is further explained below.

4.2.2.1 Forming the Appliance's Profile

Ultimate goal of any disaggregation system is to be able to identify any model of an appliance, although the different models might present fairly different characteristics, such as the power level during their operation. In an attempt to create a generic profile of the fridge appliance, the Profiler module finds the range of values that the power consumed by the various fridge models fall in. The second characteristic of the profile is the duration that a single operational cycle of a fridge lasts, for which another range of values is formed. In order for this to be achieved, the Profiler processes the various signatures using the power changes detectors and the steady states detectors. The minimum and maximum values observed in the signatures form the fridge's profile. The output of the Profiler looks like the following table:

Table 2: Example of Profiler's Output

	Minimum Value	Maximum Value
Power Level	82 W	511 W
Duration of Operation	280 sec	4258 sec

4.2.2.2 Finding Relevant Power Changes, Forming the Candidates

After the profile of the appliance has been formed it is given as input to the Disaggregation module, which is then able to process an aggregate signal. The first step is the detection of power changes in the aggregate with a magnitude that falls in the range provided by the appliance's profile. Positive changes with such magnitude are possible turn ON events of a fridge, negative changes with such magnitude are the possible turn OFF events of a fridge. By detecting the whole range of magnitudes any model of fridge can be detected, smaller or bigger, as long as its consumption falls within the range of the previously profiled models. This step is presented in Figure 4.6a.

The process continues by combining the found positive changes with the negative changes in what may be a pair of turn ON and OFF events of a fridge cycle. A candidate is created out of each combination of a positive change with every negative change in its proximity. Since the combination of the two changes will form a candidate fridge cycle, the distance of the two changes must be within the limits given by the appliance's profile for the duration of a fridge's cycle. Various candidates created during this step can be seen in Figure 4.6b.

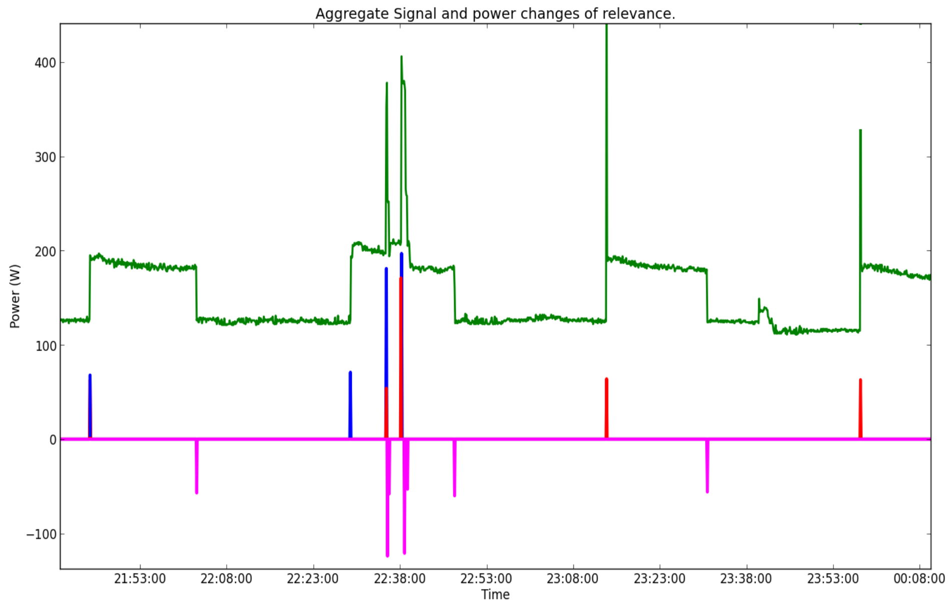
4.2.2.3 Finding the particular model's characteristics

A fridge turns ON and OFF in regular intervals, even throughout the night. In addition, by comparing several signatures of the same model it has been observed that the power consumed (and respectively the magnitude of the power changes at the turn ON and OFF events), the duration of each operation, as well as the interval between two consecutive operations are fairly stable for most of the models (see Figure 4.9). Based on these observations, the Disaggregation module tries to infer the characteristics of the particular fridge of the house by extracting the most common value for these features met in the various candidates. If this process is done over a long enough period of time, for instance aggregate data from a whole day, it can be inferred with confidence that most of the candidates are the true fridge operations, thus the values extracted are accurate. This can be further supported by the fact that the fridge also operates during the night, when not many other appliances operate. Thus many of the candidates are the true operations of the fridge during the night, increasing the accuracy of the extracted values.

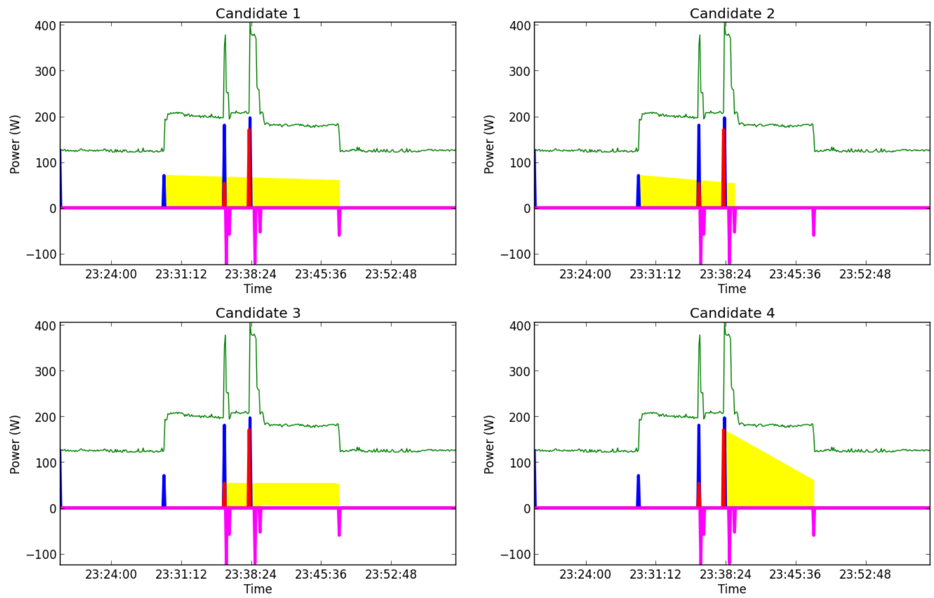
4.2.2.4 Assigning Scores

In order to evaluate the confidence of a candidate being a true instance of a fridge operation, the candidates are assigned scores. Four scores are assigned to each candidate with a value from zero to one:

- Magnitude of its positive change in relation to the most common value found among the candidates.



(a)



(b)

Figure 4.6: (a) Aggregate signal and the detected relevant power changes. Both instant positive changes (blue) and spike changes (red) are detected. (b) Four candidates created by combinations of the positive and negative detected changes.

- Magnitude of its negative change in relation to the most common value found among the candidates.
- Its duration in relation to the most common value found among the candidates.
- Its distance from its three previous and three next candidates, in relation to the most common value found for the distance between consecutive candidates.

4.2.2.5 Disaggregation

The last step, the actual disaggregation is achieved with a technique inspired by the 'Sliding Window' technique used in Photo Optical Character Recognition. A window of width 150% of the found most-common-candidate-duration swipes the data and stops at the first candidate with an overall score over 50%. In this section of the data it chooses and finally disaggregates the candidate with the highest confidence. This last step is required since more than one candidates may be formed by the same positive change combined with different nearby negative changes, with more than one of them having a fairly high score. After the top candidate has been selected, the window continues swiping the rest of the data.

4.3 Evaluation

The above technique has been tested on the data from house-1 and house-6 of the REDD dataset. For both cases, the test covered aggregate data for approximately 2.3 days period (200k seconds). The results are presented in Table 3.

Table 3: Evaluation of Fridge disaggregation

	REDD: House-1	REDD: House-6
Actual Number of Signatures	49	45
Actual Power Consumption (excl. standby power)	2.78 kWh	4.54 kWh
Number of Signatures Disaggregated	41	37
Number of True Positives	41	34
Number of False Positives	0	3
Power Consumption Correctly Disaggregated	1.92 kWh	2.99 kWh

In the case of house-1, the technique disaggregated accurately 83% of the signatures, with 69% accurately disaggregated energy consumption. In this particular house, however, the main that the fridge is connected to is a fairly tranquil one. In order to get an overall better idea of the accuracy, the technique was also tested on data from house-6 of the same dataset. In the case of house-6 the percentage of accurate signature disaggregation was 75% with consumption disaggregated equal to 66% of the actual value. In this second case, false positives were also detected, reducing the detection accuracy to 91.9% (Number of True Positives / Number of Disaggregated Events, as defined in [22]). The disaggregation of the refrigerator has been attempted in previous work with rather successful results (for instance [22] and [19]). As it was not the main aim of this research, no further attempts were made to increase the accuracy of the technique.

4.4 Conclusions

The disaggregation of the fridge was mostly based on its frequent operation throughout the day, in fairly stable time intervals. Other appliances do not present this characteristic. However during the development of this prototype certain parts of the design were found promising and were adopted in the design of the final system.

The Profiler module designed and implemented in this version was found to be an interesting step towards creating a generalized signature of an appliance. If designed and extended in an appropriate manner it could provide a profile able to characterize the various models of a certain appliance type. It can also be, however, one of the main reason of disaggregation failure if a new model differs significantly from the ones processed during the profiling phase.

The method of forming the various candidates in the aggregate out of combinations of the relevant power changes was also fairly successful. It was observed, however, that during periods that an appliance creates many power changes in the aggregate, a big number of candidates is created. A method to limit the number of candidates being evaluated should be found, in order to avoid computational inefficiency.

Using scores in order to evaluate the confidence of a candidate presented fairly positive results. The soft-limit of 50% confidence used in this design however would not suffice for more complicated appliances. In order to avoid the detection of false positives the use of a classifier should be adopted.

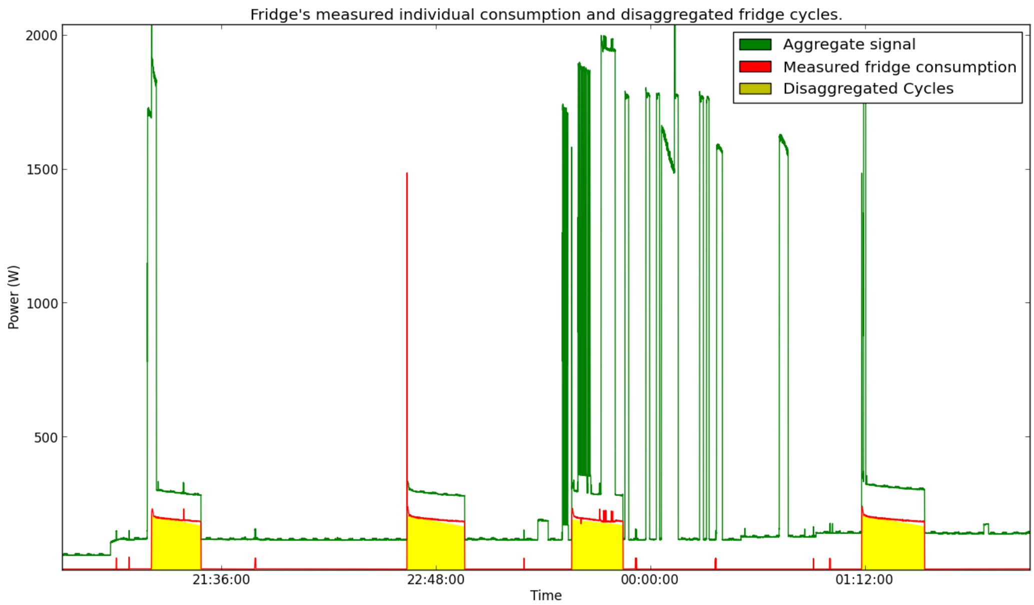


Figure 4.7: Examples of accurate disaggregation.

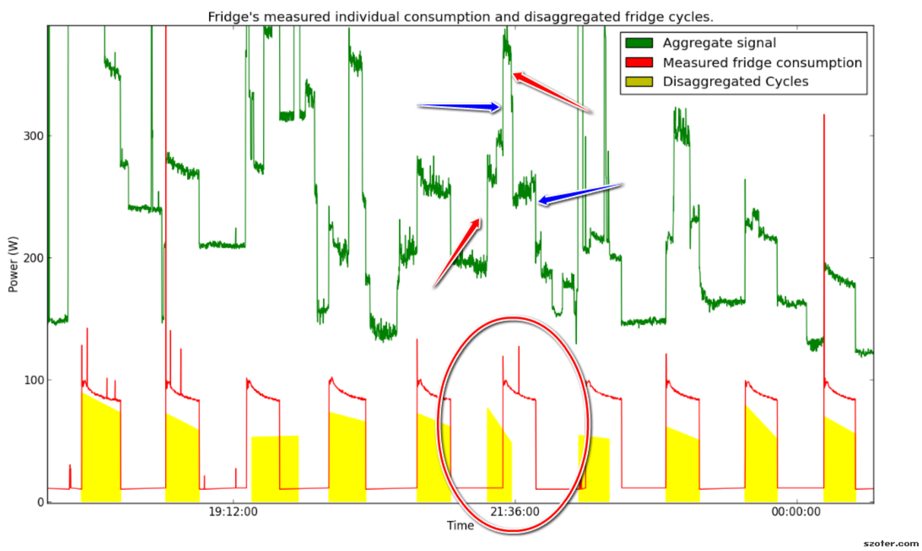
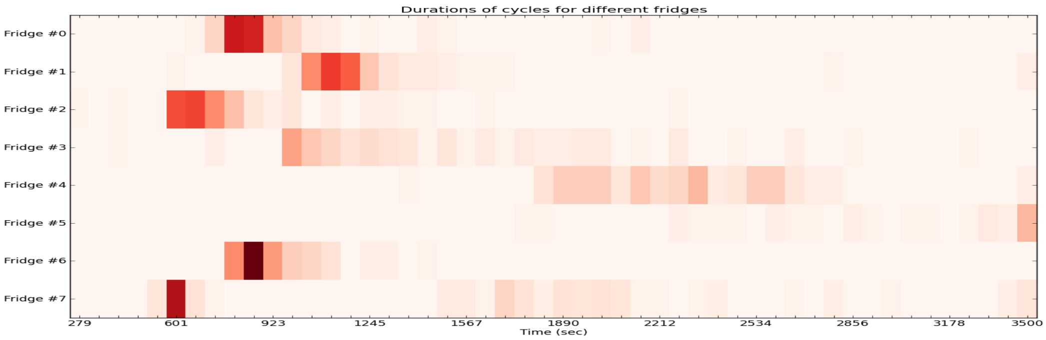
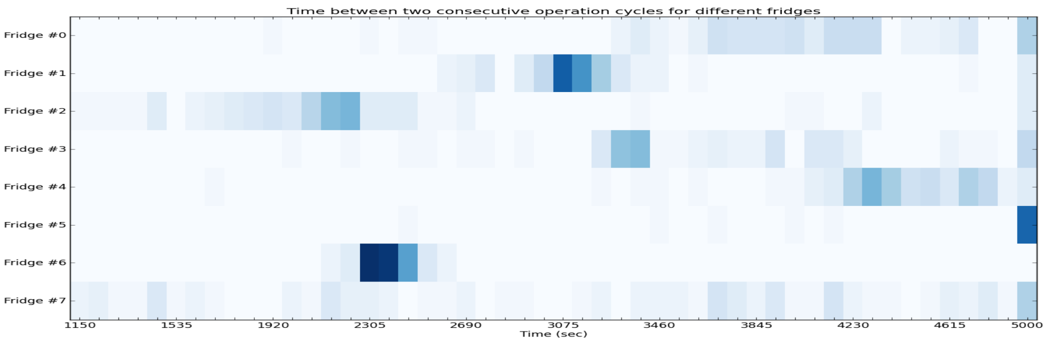


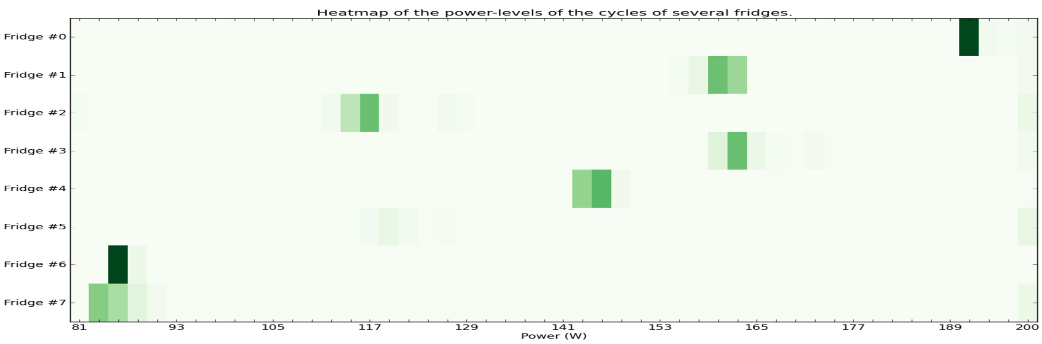
Figure 4.8: Example of false disaggregation during development of the system. The blue arrows point out the power changes in the aggregate signal created by the fridge. A wrong candidate is disaggregated, created by the power changes indicated by the red arrows.



(a)



(b)



(c)

Figure 4.9: Different models of fridges and the power (a), duration of an operation (b) and time interval between two consecutive operations (c), as observed in their signatures in the available datasets. Dark colours indicate frequent occurrences of these values.

5 Forming a Generic Profile of a Multi-State Appliance

The initial inspiration towards a methodology for forming a profile of an appliance was given by methods used in Optical Character Recognition (see section 2.3). In order to recognize a character, various feature-extraction algorithms were developped in order to evaluate its shape. For instance, it may be broken down into lines, curves and loops. Capital 'A' for instance consists of three lines, whereas 'D' can be broken down in a vertical line and a curve. The relative positions of these parts are also extracted. The result of this process, a 'map' of the character which is invariant of the size or font, is compared with a known dataset with profiles of all the known characters of the alphabet. If a close match is found, the character is classified.

Multi-state appliances like the dishwasher, the washing machine or the dryer have a rather complicated operation cycle. Methodologies that are successful for ON/OFF appliances, based for instance on detecting and clustering positive and negative changes of similar magnitude fail in accuracy when tried upon multi-state appliances. Moreover, the appliances stated above usually operate at fairly similar power levels (see Figure 3.7), making it hard to determine by which appliance a power change was created when processing magnitudes alone. The first observations presented in section 3.2.1 (also see figures 3.4, 3.5 and 3.6) suggested that an approach based on a larger number of features could be more successful.

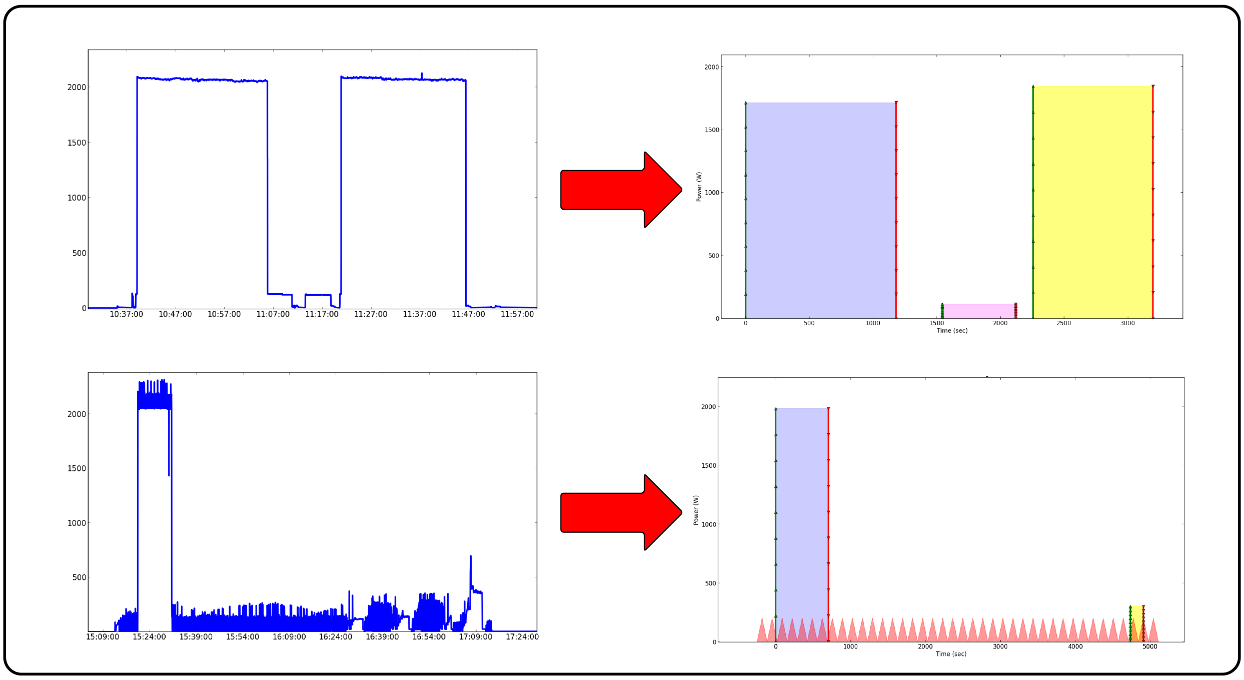


Figure 5.1: The main parts that consist the signatures of the appliances can be extracted and mapped, in order to determine the rough shape of the appliance's signature. Complemented with additional features observed in the signatures of the appliance, a generic profile of the appliance can be formed.

The methodology developed initially aims at identifying the main parts that consist the signatures of an appliance. By extracting the parts that are consistently observed within the signatures of the different models of an appliance, it is possible to create a parameterized 'map' of the appliance's signature, which roughly describes the shape of the signatures. Complemented with additional features extracted from the signatures, the generic profile of an appliance is formed. This profile essentially describes an appliance as a 'set of features', the combination of which may allow a distinction of the appliance within the aggregate signal.

5.1 Design of the System

In this section is presented an overview of the system designed in order to enable the processing of the signatures and the extraction and analysis of features. From these features a generic profile of an appliance is formed and the methodology is tested by attempting the actual detection and disaggregation in the aggregate signal. A flowchart of the whole system is presented in figure 5.2. The various components of the system are described in more detail in the corresponding sections of the report. The design of this system was also influenced by the Prototype system described in the previous section, a reading of which could provide a better intuition to the reader.

5.2 The Profiler

Aiming towards the main objective of forming a generic profile of an appliance, the Profiler module was designed. Inspired by feature extraction used in OCR, it processes the signatures of the various models and through certain steps tries to determine the rough shape of the signatures by identifying and 'mapping' parts of the signature that are consistently observed in the signatures of the various models. After the main parts are identified, the mapping is achieved by evaluating the magnitudes of the main power changes in the signature and their relative positions. The Profiler afterwards evaluates the existence of features in the signatures, like the existence of spikes, rapid changes and parts of the signatures that are repeated during an operation of the appliance. The output is the profile of the appliance, a collection of features that describe the generic shape of the appliance's signatures.

The above process is preceded by extracting from each individual signature the main steady states and the power changes that lead from one main steady state to the other, process done by the secondary module Signature Mapper.

5.3 The Pattern Recognition Module

In order to test the efficiency of the approach it was needed to develop a module that would detect possible operation cycles of the appliance in the aggregate signal and compare them with the generated profile. The Pattern Recognition module receives as input the profile of an appliance, as formed by the Profiler, as well as the data to be scanned. Making use of the profile's parameters, it detects in the aggregate signal patterns that are possible operations of the appliance (candidates). It then proceeds to compare them with the profile and evaluate them, assigning scores in different areas. This results in a list of scores for each candidate, which is provided to the classifier for the final step, the classification.

The Pattern Recognition module also plays a part in the training phase of the classifier. For the training phase, the module receives as input a set of appliance signatures, labelled as positives or negatives depending on whether they belong to the actual appliance or another type. In a similar fashion as described above, it detects the candidates in the signatures and evaluates them. The high scored candidates from the positive signatures and the low scored candidates from the negative signatures are passed to the classifier for training.

5.4 The Classifier

The final part of the system is a One-vs-All Classifier which is called to make the final decision on whether a candidate is a true operation of the appliance or not. As input, the Classifier receives scores provided by the Pattern Recognition module, after evaluation of the candidates. These various scores are actually the features for the classifier. Initially the classifier needs to get trained on the labelled as positive or negative scores. The training needs to be done once. After that, the

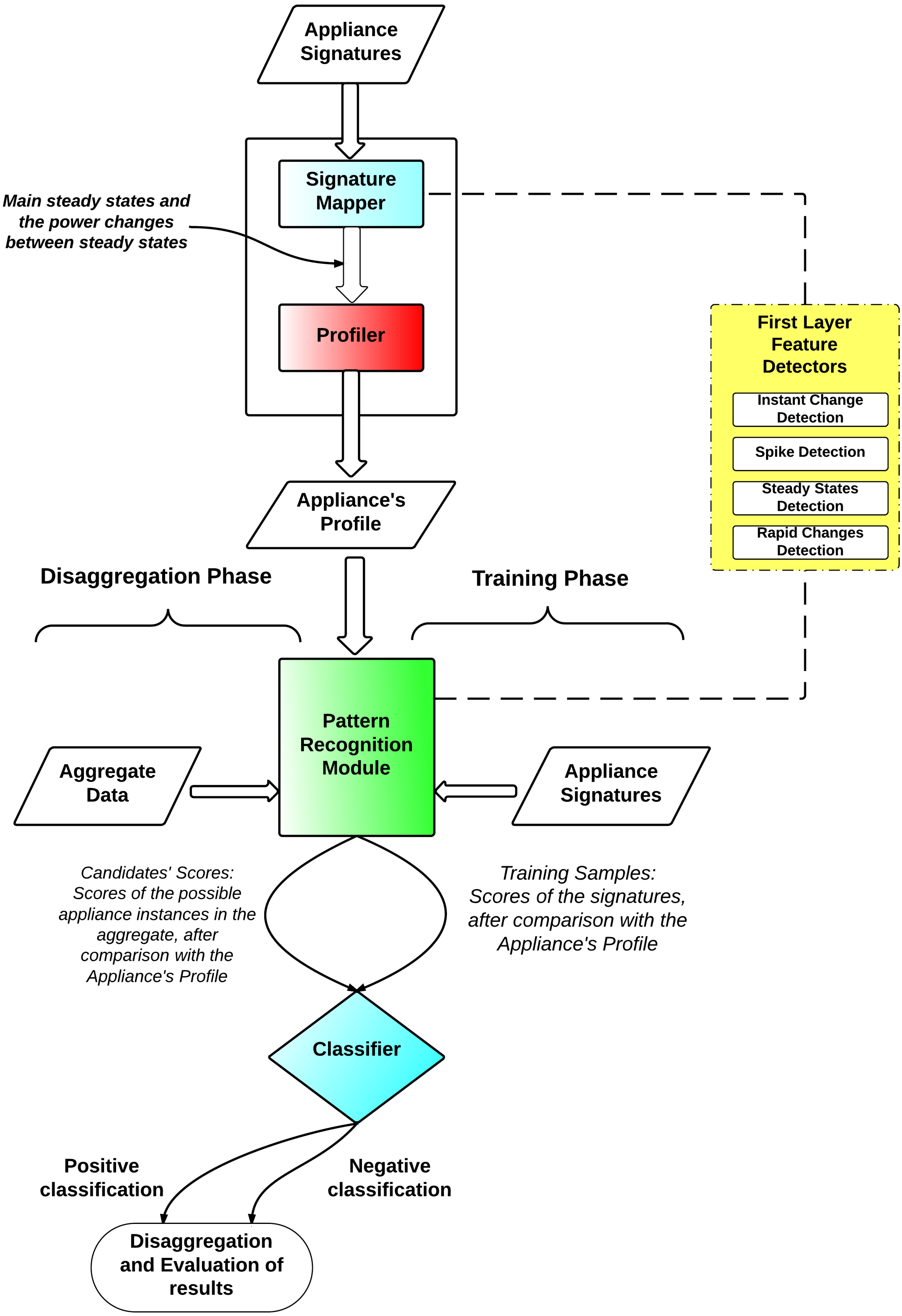


Figure 5.2: Flowchart of the system's design.

classifier 'knows' what scores correspond to the true appliance and is ready for the disaggregation phase. During the disaggregation the classifier receives a list of scores for each candidate found in the aggregate data by the Pattern Recognition module. It then proceeds to compare them with the trained scores and classifies a candidate as a true operation of the appliance if the scores of the candidate are close enough.

5.5 Why not a Sample-by-Sample Pattern Matching Approach?

Certain techniques in character recognition and other areas of computer vision follow a pattern matching approach. The to-be identified image is compared to a collection of stored images on a pixel-by-pixel basis. Alternatively but in a similar fashion, a classifier is trained on a collection of images, where each pixel represents a feature for the classifier. The classifier then processes pixel-by-pixel the new image and classifies it positively or negatively depending on how close its pixels are to the trained set.

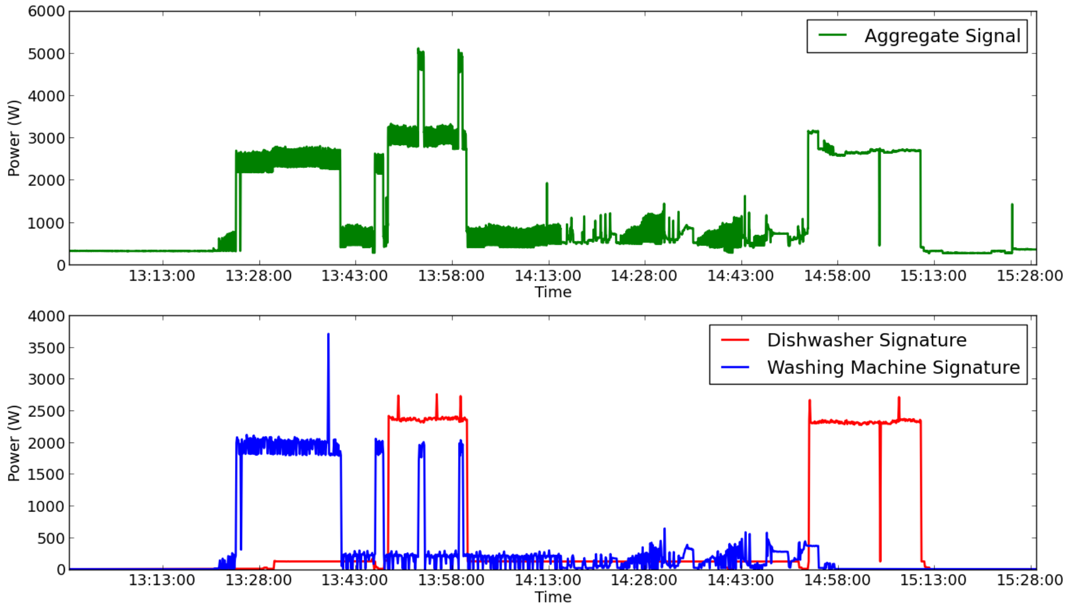


Figure 5.3: A dishwasher and a washing machine operating at the same time. A pattern matching technique would fail recognizing their 'skewed image' in the aggregate signal, generated due to the additive nature of power consumption.

In the case of energy disaggregation, an appliance's signature could be considered as a 'one dimensional image' with its timestamp representing a pixel and the corresponding power value representing the color of the pixel. This way, it could be treated with a similar technique with the ones described above? Unfortunately the problem is more complicated and such an approach is likely to fail.

In order to detect a character or another object in an image (for instance a face), the computer vision techniques mentioned above process images where the object is clearly visible. In the case of energy disaggregation however, the aggregate signal is produced by the sum of the power consumption signals of all the appliances. If two appliances operate at the same time, their signals are mixed, thus their 'image' in the aggregate signal is skewed (see Figure 5.3). An analogy with character recognition would be two characters written on top of each other. A pattern recognition system, like the suggested one, may manage to extract enough features in the 'image' that indicate the operation of a certain device.

6 Profiling an Appliance

The various models of an appliance present several differences in their operation. For instance a large model of a dishwasher consumes higher amount of power than smaller models. Differences can be observed even between two operations of the same model when a different program is selected. As an example, many models of washing machines allow the user to choose the number of rinsing cycles. This complexity makes it difficult for a disaggregation system to detect accurately different models than the ones trained upon, when based only on a few features such as power consumed and duration of use.

Despite this complexity, the initial investigation of the appliances' signatures, which was presented in section 3.2.1, revealed that some patterns can be consistently observed in the signatures of an appliance. For instance the dishwasher's heating element usually turns up twice during its operation. Other elements of the dishwasher consume a smaller amount of power throughout the operation. The washing machine's heating element is preceded by the motor, which rotates for almost the whole duration of the program. At least a rinsing cycle is observed near the end of the wash in any program. Finally, the dryer's heating element usually turns on and off several times, staying on for similar amounts of time in order to keep the temperature at steady high levels. These observations have been presented in 3.4, 3.5 and 3.6).

In this section are presented the steps followed in order to form a profile of an appliance. The creation of the profile is based on detecting certain parts of the signatures that are consistently present in the signatures of the various models. It can be perceived as an attempt to describe the approximate shape of the signatures with the main 'blocks' observed in their shape. The profile describes the values that the various characteristics of the signatures take for the different models, for instance the magnitudes of the main power changes and the distances between them. It also describes the existence in the signatures of features such as spikes, rapid changes and repeating parts of the signatures, as well as how often these features are observed. By including this information, the profile fits in a generic way all the different models of the appliance type processed during the forming of the profile. Thus it can be used by a pattern recognition system in order to detect in the aggregate patterns that are close to what the profile describes. By incorporating various features it allows for detection of the targeted appliance even if some of them are not detectable in the aggregate signal, for instance by the operation of another appliance. The various steps followed to form this generic profile are presented in the rest of this section.

6.1 Mapping each Individual Signature

When an appliance operates in a steady state, its power consumption remains at a fairly steady level. When a transition between states occurs, a power change is observed in the signature. The first step is to process each signature individually and extract the steady states and power changes that determine the main shape of the signature. This is accomplished by the secondary module *Signature Mapper*, which processes each signature individually and extracts the main steady states observed. Since the goal is to extract the rough shape of the signatures, neighbouring steady states of same power level are grouped together, discarding any short breaks in between (see figure 6.2). After the main steady states of each signature have been determined, the power changes that lead from one main state to the other are detected. The Mapper also detects the type of a power change, whether it is an instant change or a spike. This process is done for every signature provided in the data and the extracted steady states and power changes are provided to the Profiler.

6.2 Choosing the Main Blocks of the Signatures

After the profiler receives the extracted steady states and power changes of each signature from the Mapper, it processes them all together in order to extract the parts of the signatures that are

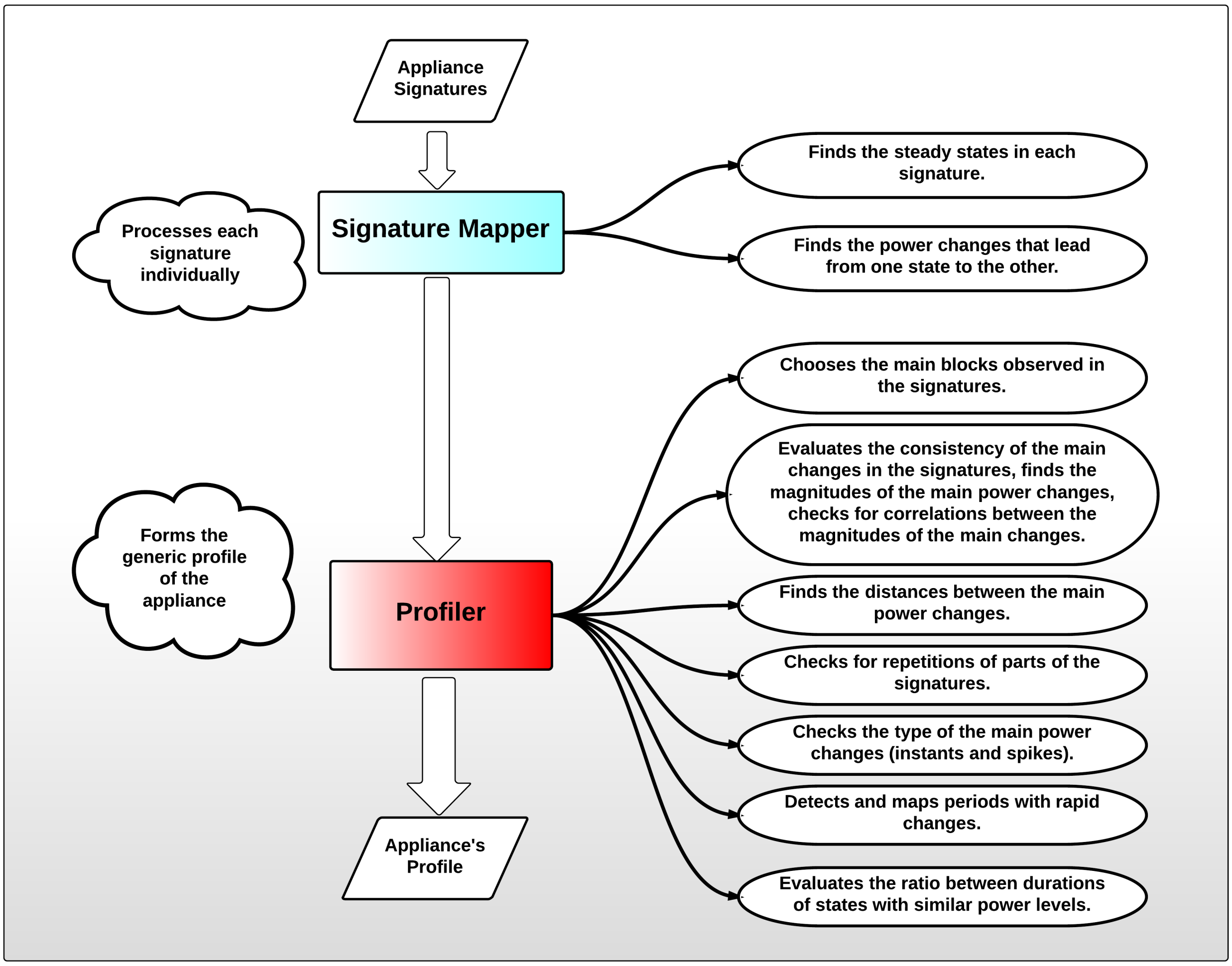


Figure 6.1: The steps towards forming the generic profile of an appliance. The Mapper processes initially each signature individually, extracting their main steady states and power changes. Then the profile is created in certain steps, by determining the consistent parts of the signatures and evaluating the existence of various features.

consistently observed in the majority of them. The area below a main steady state in a signature can be described as a 'block', because of its shape, with height equal to the steady state's power level and width equal to the steady state's duration. A block consistently observed in the signatures can be interpreted as a part of the operation that a certain element of the appliance has turned on and operates, for instance the heating element in the beginning of the washing machine's operation. The main steady states detected in the various signatures are processed in order to choose the 'main blocks' that will roughly describe the generic shape of the signatures. This process is done according to the following algorithm:

- The main steady states observed in each appliance are separated in three groups. The highest-power group with power-level between 100%-70% of the maximum power level in the signature, the medium-power group with 70%-35% of the maximum power level and, finally, the low-power group with the steady states consuming power between 35%-0% of the maximum power level. This separation can be seen in figure B.1.
- Choice of first main-block: Choose the first (leftmost) block of the highest power group, as long as one exists in the majority of the signatures. Else, choose one from the next power group.
- Choice of second main-block: Choose a block from the next power group, before or after the first chosen block, depending on where a block is found more consistently in the signatures

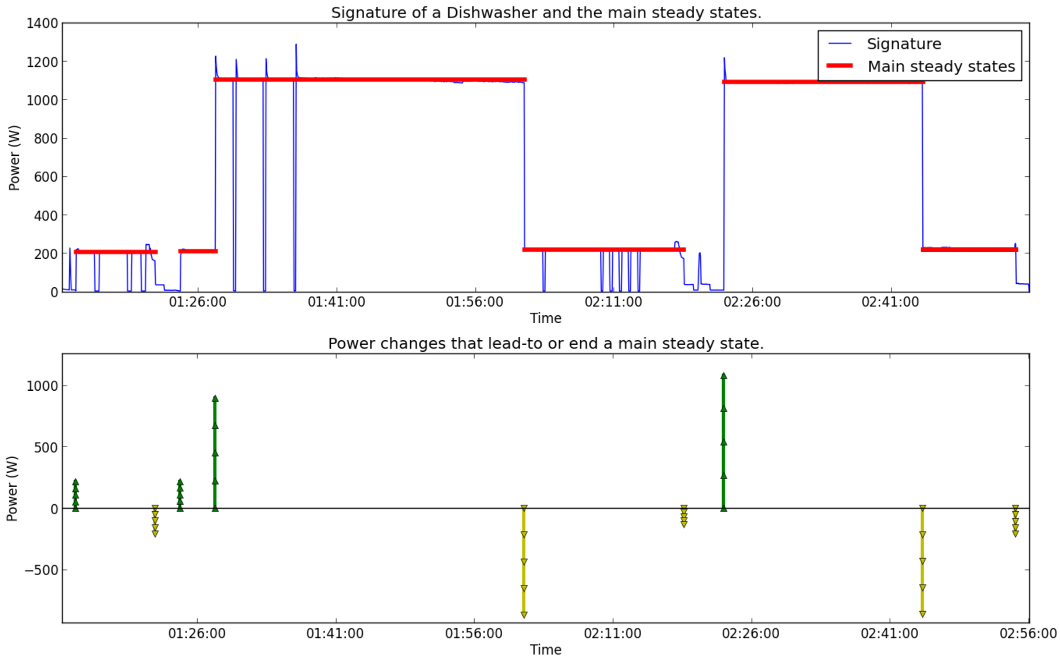


Figure 6.2: The Signature Mapper module processes each signature individually and detects the main steady states and the power changes that lead from one state to the other. Small transitions or short breaks are of little interest and are discarded, as the goal is to extract the rough shape of the signature. The type of the changes is also detected; whether they are instant changes or have the form of a spike.

and as long as one is found in the majority. If before, choose the leftmost block. If after, choose the rightmost block. If no case is consistent, try choosing from next power group.

- Choice of third main-block: Choose a block from the next power group, before, between or after the two previously chosen blocks, as long as one exists in the area for at least 50% of the signatures. If no case is consistent, check if choosing the second main-block from this group and the third main-block from the previous power group leads to consistency. If not, do not choose third block.
- In the case that a power group is empty (no steady states belong to it for most signatures), the next block can be chosen again from the highest power state.

This algorithm might seem quite unintuitive. The result however is simply the choice of up to three blocks from the different power-groups, in certain relative positioning, consistently observed in the majority of the signatures. Having determined these consistent blocks allows the 'mapping' of the signatures' shape in relation to them, as presented in the next steps. Examples of the chosen blocks resulting by the algorithm in the case of the dishwasher, the washing machine and the dryer are visualized in figures 6.10a, 6.10b and 6.10c.

Positioning of the chosen blocks and the steady states they correspond to

The main blocks represent the main parts of the appliance's signatures, or at least up to three of them. Depending on the area of the signature that the blocks were chosen from, they refer to certain steady states in the signatures. The steady states that the chosen blocks refer to for the different cases of the algorithm are presented in figure 6.3. For instance the first and second block for the dishwasher (see Figure 6.10a) have been chosen from the same power-group of steady states. However the first block corresponds to the first steady state of the group, whereas the second block corresponds to the last one of the group. These certain steady states, their durations,

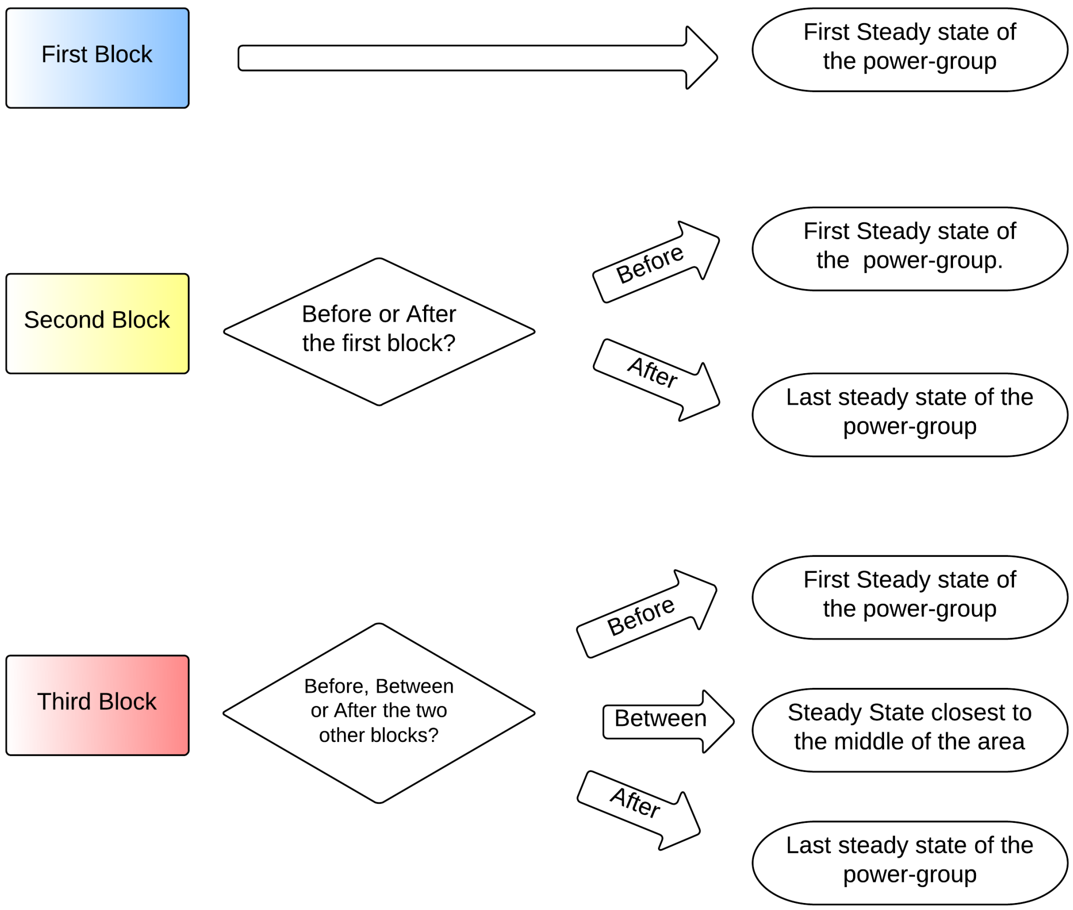


Figure 6.3: The chosen blocks from the algorithm described in subsection 6.2 refer to certain steady states in the signatures. The steady states they refer to depend on the power-group they were chosen from and the position that the block was chosen, in relation to the other blocks. These steady states are the ones that the appliance's profile is built around.

the magnitudes of the positive and negative corresponding power changes and the distances from each other are the characteristics that will help form the 'map' of the signature in the next steps.

Design of the algorithm

The 'blocks' chosen in this subsection are the basic patterns around which the profile of an appliance is created in the next steps. They are also the patterns that are going to be looked for in the aggregate during the pattern recognition stage. The aim of the above algorithm is to choose at least one block from every power-group of the steady states, in order to represent them in the generic profile of the signature generated afterwards. This design does not immediately limit the capabilities of the technique, although initially it seems to leave out big part of the signature. Since a block is chosen from every power-group, it is possible to add to the model other parts of the signature with similar power-level as 'repeated parts' of the signature, as long as they are consistently observed in the signatures. This process will be further explained in the corresponding section.

Since the aim is, in a sense, to find all the consistent parts in the signatures and roughly describe its shape, one may ask why to choose only three blocks. The answer is that these chosen blocks represent the basic pattern that the Pattern Recognition module will try to detect in the aggregate signal. To be more precise, what is being detected during the pattern recognition phase are the power changes in the aggregate signal with magnitudes similar to the power changes that make up the edges of the main blocks. By any combination of such power changes detected, a candidate (possible operation of the appliance) will be formed and evaluated. The use of a forth block increases exponentially the number of combinations, making the disaggregation particularly inefficient in some cases that many such power changes occur. This will be further understood after the analysis of the Pattern Recognition module.

6.3 Main-Changes of the Profile

As explained above, the profile should describe the generic shape of the appliance's signature, in order for a pattern recognition system to 'know' what kind of patterns to look for in the aggregate. At this point it is important to note that in the aggregate signal the steady states in power are of little use because of the additive nature of the power signal. For example the steady state of the heating element of the dishwasher might not be visible because of an other appliance turning on during its duration. What is actually observable are the power changes created when an appliance changes state.

For that reason, the profile of an appliance is formed around the *main power changes*. Main power changes of an appliance are the positive and negative changes that make up the edges of the chosen main blocks of the profile. The main changes of the profile of the dishwasher, the washing machine and the dryer can be seen in figures 6.10. The profile of the appliance is formed around them in the sense that it incorporates characteristics such as the magnitudes that these main changes take in the signatures of the various models. They are also used as the reference for 'mapping' the signature of the appliance, for instance for defining the area that the rapid power changes begin when a washing machine turns on.

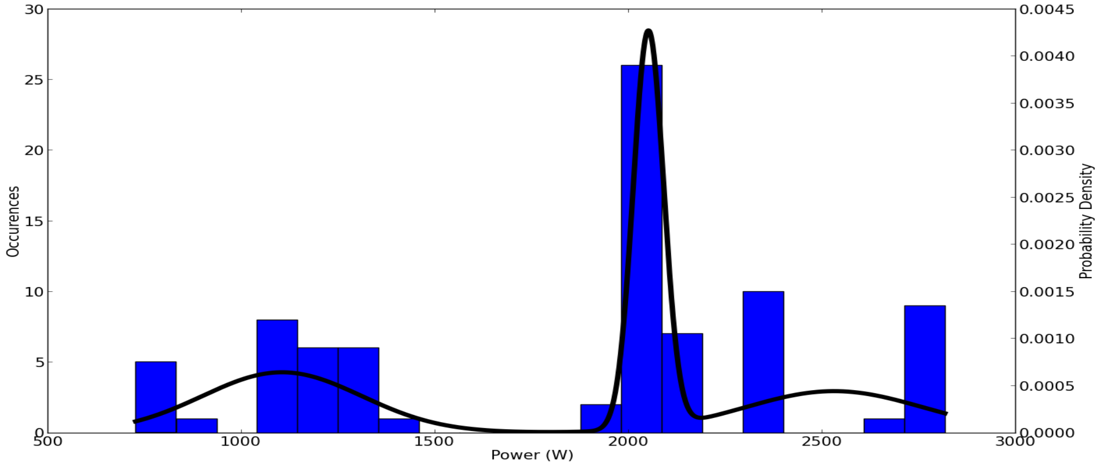
6.4 Magnitudes of Main Power Changes

After the main blocks of the generic profile have been chosen, it is possible to start forming the profile of the appliance in more detail. The next step towards this goal is to find the power that the several models of the appliance consume in the corresponding steady states, as well as the magnitudes of the relevant power changes. This way the profile is capable of describing the power changes generated during the appliance's operation, which should be looked for in the aggregate signal for disaggregation.

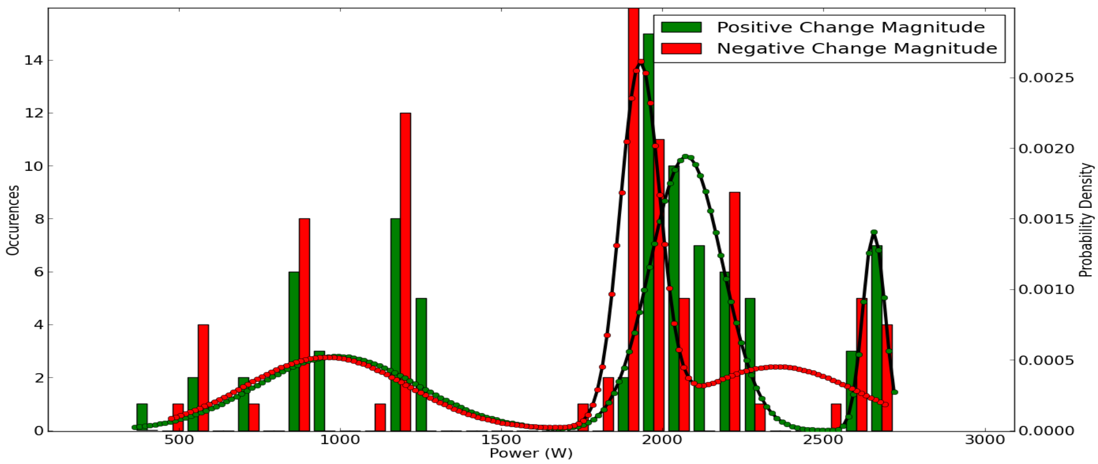
The power consumption during the various steady states and subsequently the magnitude of the corresponding power changes varies from model to model. Histograms depicting the values extracted for the first heating cycle of the dishwasher from various signatures available in the datasets are presented in figure 6.4. For the rest of the appliances the results can be found in the relevant section of the Appendix. To incorporate this information about the power consumption of the various models into the profile of the appliance, a density function is fit to the extracted values.

Initially, a Gaussian function has been used for such purpose. In some cases the values seemed to present a Normal distribution. In other cases the distribution differed significantly however. The use of the Gaussian function was attempted based on the belief that the datasets are not particularly complete, so a Gaussian could generalize over values that were missing. In later steps of the project, this choice was abandoned. Gaussian mixtures were adopted instead, which are capable of fitting such distributions with fairly abstract shapes much more accurately.

A Gaussian Mixture Model fits a density function to the samples. The Gaussian mixture is composed by a number of components, each component is itself a Gaussian function (see section 2.4 for more information on Gaussian Mixture Models). The components of the GMM are automatically determined, up to a number of three components, depending on which number better fits the distribution. The upper bound of three components was used in order to avoid overfitting the data.



(a)



(b)

Figure 6.4: (a) Power consumption during the heating cycle of the various dishwasher models in the available datasets. Up to ten signatures were processed from each model, with some models having less signatures available in the data. (b) Histogram of the magnitudes of the positive and negative power changes of the first heating cycle observed for the various dishwashers. Gaussian Mixture Models have been used in order to determine the density function that best fits the data.

This choice also agrees with the fact that there can be three sizes of appliances, small, medium and large models. Such density functions generated for the dishwasher can be seen in figure 6.4.

In figure 6.4, it is important to notice the difference between the values observed for the magnitudes of the positive and negative power changes associated with the first block of the appliance, the first heating cycle. The magnitudes found for the positive changes that lead to the steady state of the first heating cycle and the negative changes at the end of it are different than the power level of the state, and slightly different from each other. The reason is that in some signatures another steady state precedes the first heating cycle, or another one follows it. Thus, the power changes have magnitudes that correspond to the difference of the power levels of these steady states.

The density functions presented above are one of the main parameters of the profile describing the appliance. They are one of the factors that allow the pattern recognition module to estimate the probability that a power change observed in the aggregate data during the disaggregation is a power change created by a change of state of the profiled appliance. If changes of similar magnitudes have been observed frequently in the signatures of the models during the creation of the profile, the power change in the aggregate has a high probability. In a similar way, density functions will be used to describe more aspects of the profile in the next steps.

6.4.1 Inconsistent States and Power Changes

When choosing the main-blocks of an appliance's profile, it is a requirement that the corresponding states appear in the majority of the signatures. It does not mean however that they do appear in every single signature. In the case that a certain state does not appear, neither do the corresponding main changes. It could also be the case that a certain block appears in a signature, but one of the corresponding main changes does not. This case can be easily understood by observing a signature of a dryer and the blocks chosen for the profile of the dryers in figures 6.9 and 6.10c. In order to take this into account, the profile is complemented by a percentage describing the consistency of each main power change of the profile, the percentage of models' signatures that the main change actually appears.

6.5 Correlation between Magnitudes of Power Changes

The main blocks chosen to form the generic profile of an appliance correspond to states that are consistently found in the signatures of the different models. However, different models may present different intermediate states. A certain model may go from main state A to the zero power state for a while, and then to main state B. Another model may have another intermediate state between the two main states. To model such behaviour the profiler is complemented with entries describing a correlation of the magnitudes of the main power changes involved. This can be better understood by referring to Figure 6.5.

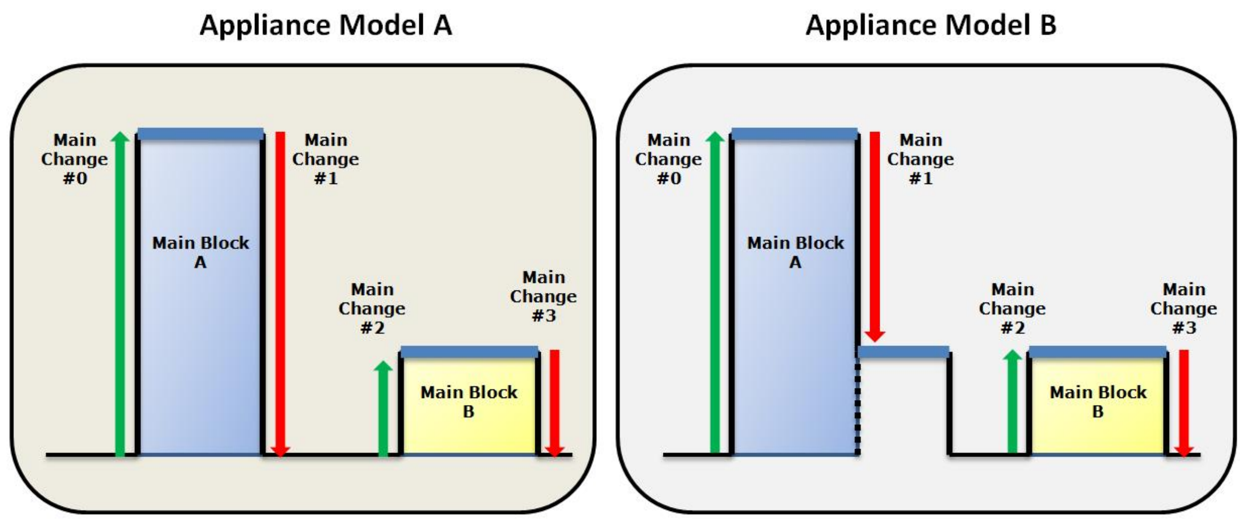


Figure 6.5: Different models may present different sequences of states. For instance a model may move from the first main state (corresponding to main-block A) to the zero power state, then move to the second main state (corresponding to main-block B). Another model may move to an intermediate state. In order to model such behaviour an entry is added to the profile, describing a correlation between the magnitudes of the main changes such as:
Magn(Main Change#0) = Magn(Main Change#1)
OR Magn(Main Change#0) = Magn(Main Change#1)+Magn(Main Change#2).

6.6 Distances between the Main Power Changes

The next important parameter of the profile is the time distance between each main power change. It allows the profile to describe the duration of the main states of an appliance, for example for how long the heating element stays on. It also allows describing the time between two different parts of the appliance's operation, such as the time from the beginning of the heating cycle to the beginning of the rinsing cycle of the washing machine. The Profiler evaluates the distances between each main-change. For each pair of main changes, the distances observed in the various signatures are extracted and a Gaussian Mixture Model is used to fit a density function to the values, in a similar fashion as described above for the magnitudes of the power changes.

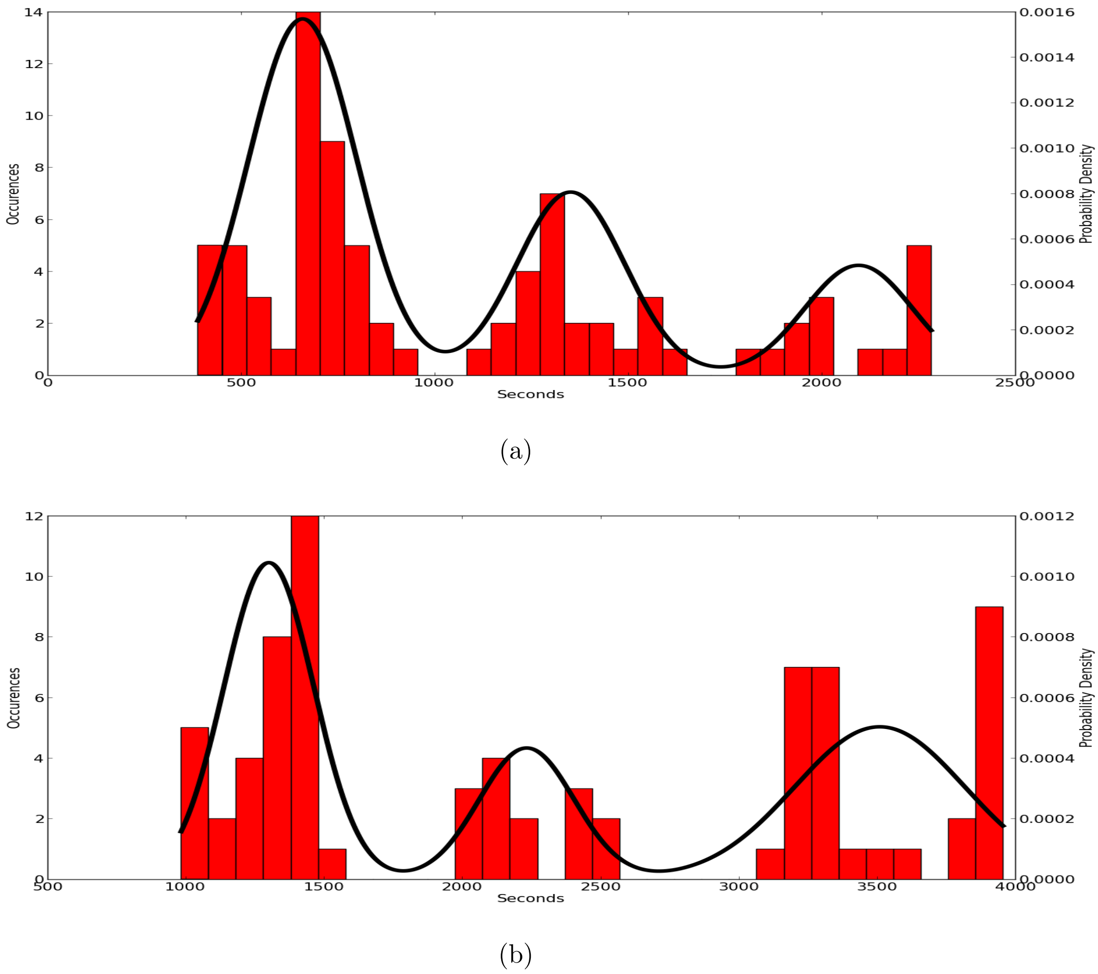


Figure 6.6: (a) Histograms and density functions of (a) the duration of the first heating cycle, (b) the time interval between the beginning of the two heating cycles of a dishwashers, as extracted from the various signatures in the available datasets.

6.7 Repeated Parts of the Signature

While investigating the data, it was noticed that during the operation of some appliances, some parts of the operation are repeated. The more distinctive such behaviour has been observed for the dryer, easily noticed in dryers' signatures. The heating element of the dryer turns on and off several times throughout its operation for most of the available models, in order to keep the temperature high and stable (see Figure B.9). It was also noticed that the duration of most of the cycles, as well as the time interval between them is fairly stable. A similar behaviour was observed for the heating element and the rinsing cycle of the washing machine, although this behaviour is less consistent between the various models than the case of the dryer (see Figure B.10).

These repeating cycles are detected by the Profiler as repeats of the main-block with similar power consumption. The signature duration is divided into areas before,between and after the main blocks, allowing the Profiler to detect the area that the repeats occur. If such repeated cycles occur consistently in an area for the majority of the signatures, the feature is added to the generic profile. The information added to the profile describes the main block that the repeating parts have similar power levels with and the area that they occur. Finally, an entry describes the number of the repeated cycles and whether they have been observed to have similar durations and steady time intervals between them.

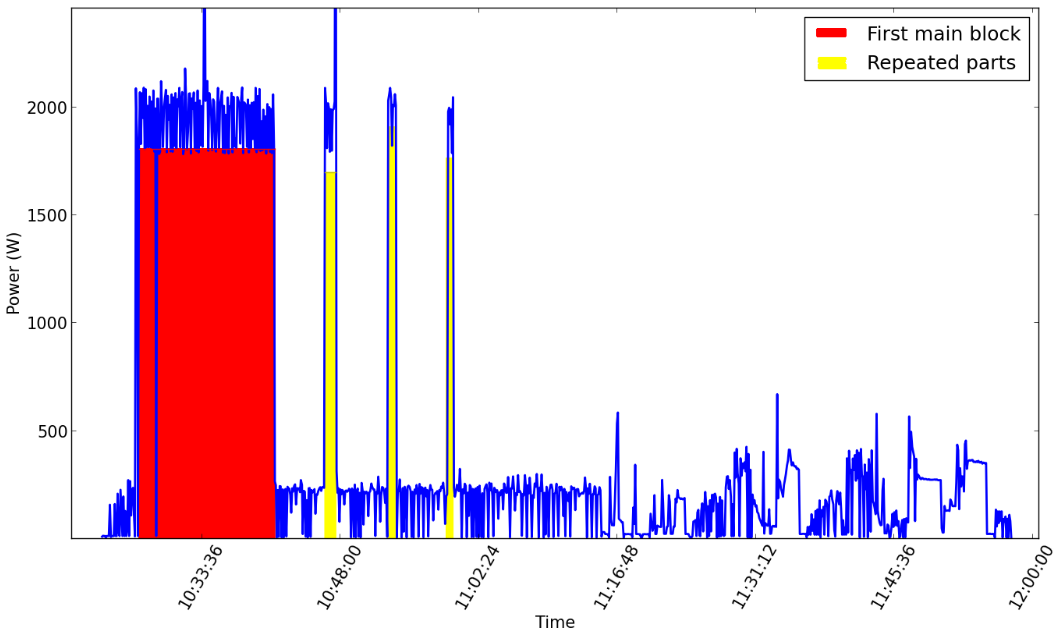


Figure 6.7: During the operation of various models of washing machines the heating element turns on more than once. The Profiler detects the subsequent heating cycles as repeats of the first chosen block, the first heating cycle.

6.8 Other features

6.8.1 Type of Power Changes - Spikes

A fairly distinctive characteristic of some appliances is the existence of positive power changes that present an initial overshoot, referred to as 'spikes' because of their shape. These spikes are mostly observed when a heating element turns on, as explained in section 4.1.2. When the power changes are detected in the signatures by the Mapper their type is also extracted, whether it is a regular change or a spike. Each main change is associated with an entry in the profile, expressing in what percent of the signatures processed the change was found to be a spike.

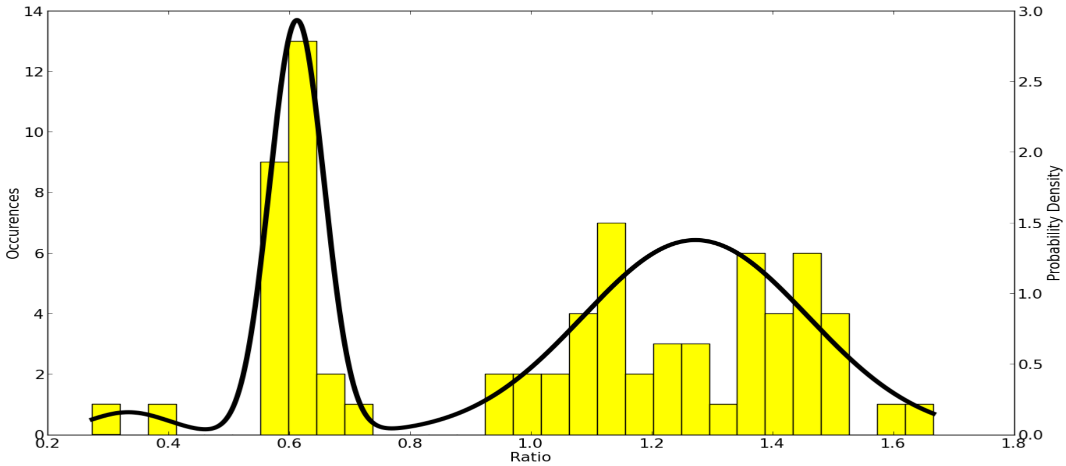
6.8.2 Rapid power changes

A very distinguishable features for certain appliances is the existence of rapid power changes at certain parts or throughout their operation. This feature can be easily observed in the signatures of the washing machine, created by the motor of the appliance. If such rapid changes are detected in the majority of an appliance's signatures, the Profiler detects the periods where they usually begin and end. These periods are mapped with relative distances from the main changes and the profile of the appliance is updated with an entry for this feature.

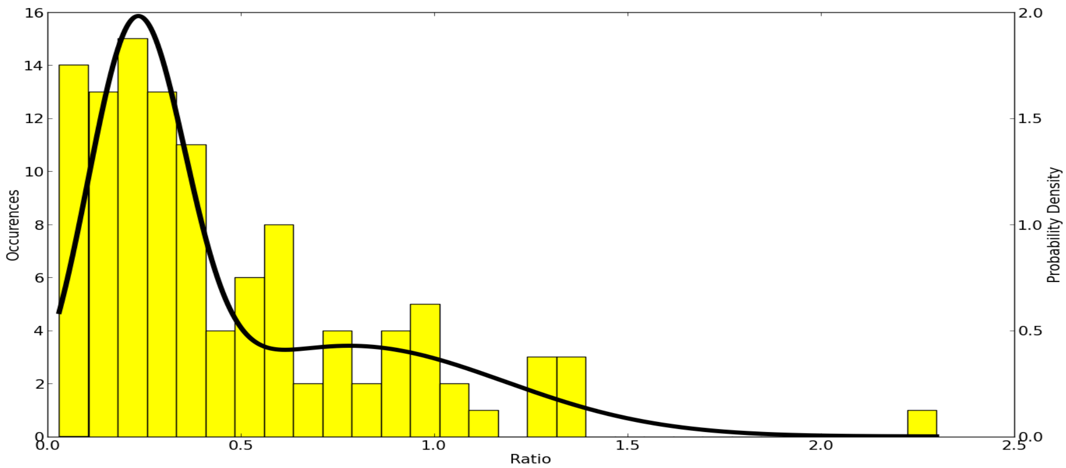
6.8.3 Ratios Between Durations of States

While investigating the appliances' signatures, it has been observed that in some cases certain parts of their operation tend to have similar durations. This can be fairly easily noticed in the signatures of the dishwashers (see 6.9a), where two cycles of the heating element tend to last for similar amounts of time. After this observation, the data were further processed in an attempt to discern more such relations. A relevant finding was extracted from the data for the dryer. In this case it was found that the subsequent heating cycles last significantly shorter than the first heating cycle of the operation. These findings are presented in figures 6.8.

In an attempt to incorporate such features into the generated profile of an appliance, the Profiler module iterates through the appliances' signatures and checks the ratio between the durations of the states that correspond to the chosen main-blocks of the profile. A Gaussian Mixture Model is used to fit a density function to the extracted ratios. Since such meaningful findings were only noticed for states of the same power level, the current version of the Profiler evaluates only these cases.



(a)



(b)

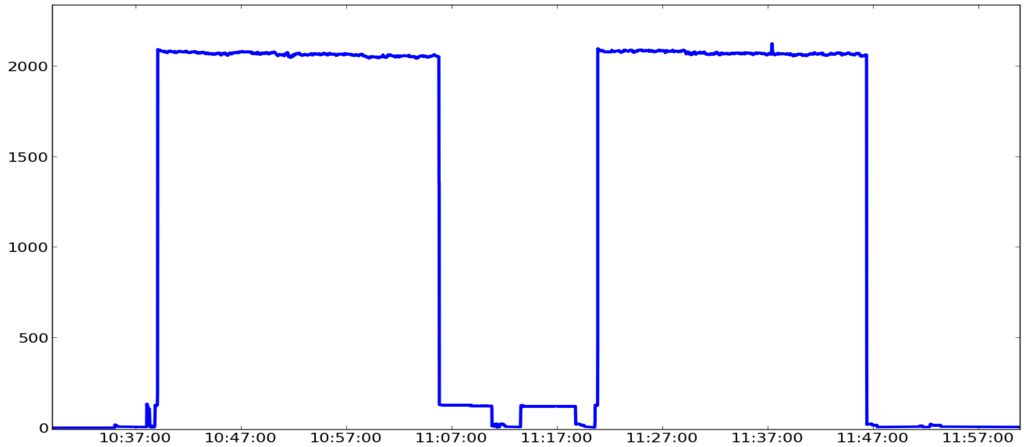
Figure 6.8: (a) Histogram of the ratios of the second heating cycle's duration to the first heating cycle's duration, for the various dishwasher models. (b) Ratios of the duration of a following heating cycle to the duration of the first heating cycle, as observed in signatures of dryers. Up to three subsequent cycles were checked for each signature. Gaussian Mixture Models are used to fit a density to the values and incorporate them to the profile of the appliances.

6.9 Why the big number of features

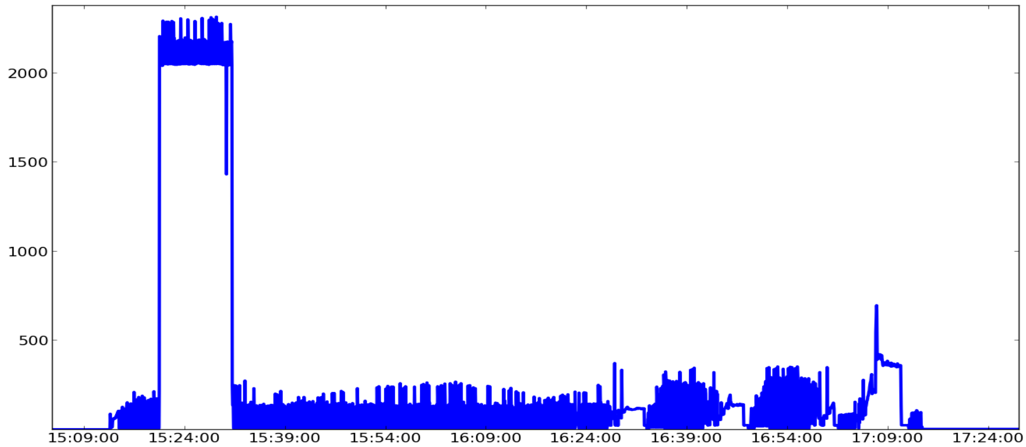
One may wonder why incorporate all these different features in the profile. Some of them may seem to describe essentially the same thing. For instance one may question adding features like ratios of durations between two states, when the distances between the main changes have already been profiled. The reason is to attempt and reduce the probability of a false positive during the pattern recognition phase. In order to understand this, one must consider that an objective of the profile is generalization. Thus it cannot be too restrictive about the values allowed, for instance for the distances between the power changes. One has then to consider that the aggregate signal is created by a big number of appliances, turning ON and OFF at random time intervals. The numerous power changes generated can be combined in several ways and create a vast amount of patterns and shapes. The generalization and the numerous power changes in the aggregate may lead to the detection of patterns during the disaggregation phase that are falsely considered as fairly confident candidates, even though a certain feature of theirs is rather off. For instance, a candidate might have a certain state duration rather long, but still acquire a fairly high probability by the corresponding GMM density function because of generalization. In that case the evaluation of another feature, for instance the ratio of durations between two states, might bring the confidence of the candidate down, eliminating a false positive. By using features that describe from different aspects the generic shape of a signature, the probability of a false positive is reduced.

6.10 Limitations Introduced

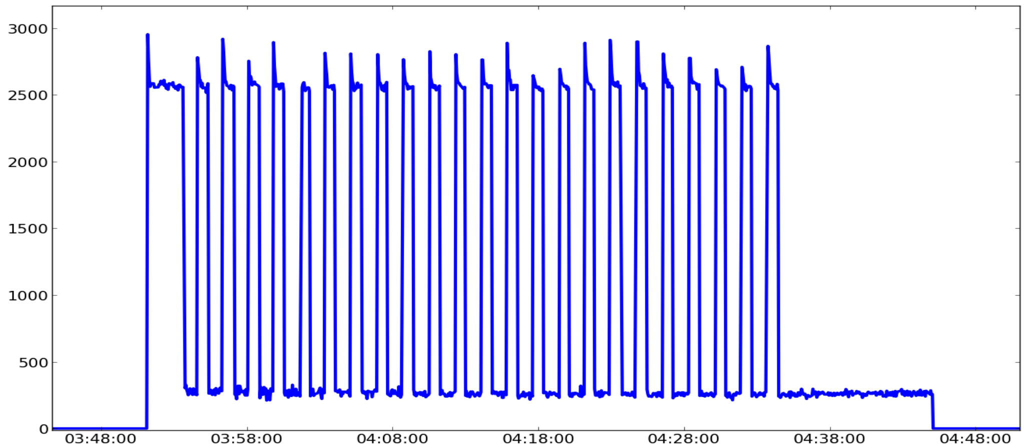
The technique that was presented forms a profile for an appliance after detecting parts of the appliance's signature that are consistently observed in the signatures of the different models processed. Parts of the signatures that are not observed in the majority of the signatures will not be chosen as a main block. Similarly, other parts of the signature, which consume similar power to a main block, will be modelled as 'repeats' of a main block only if they are consistently observed. This design was made in order for the profile to be generic and describe the whole appliance type, independent of the model. Unfortunately it does mean that parts of the signatures are left out. Parts of the signature that are not described by the profile are essentially 'unknown' to the Pattern Recognition module and will not be detected. However, taking into account that those parts are not observed in many appliances, the overall efficiency of the technique is not expected to be greatly reduced.



(a)

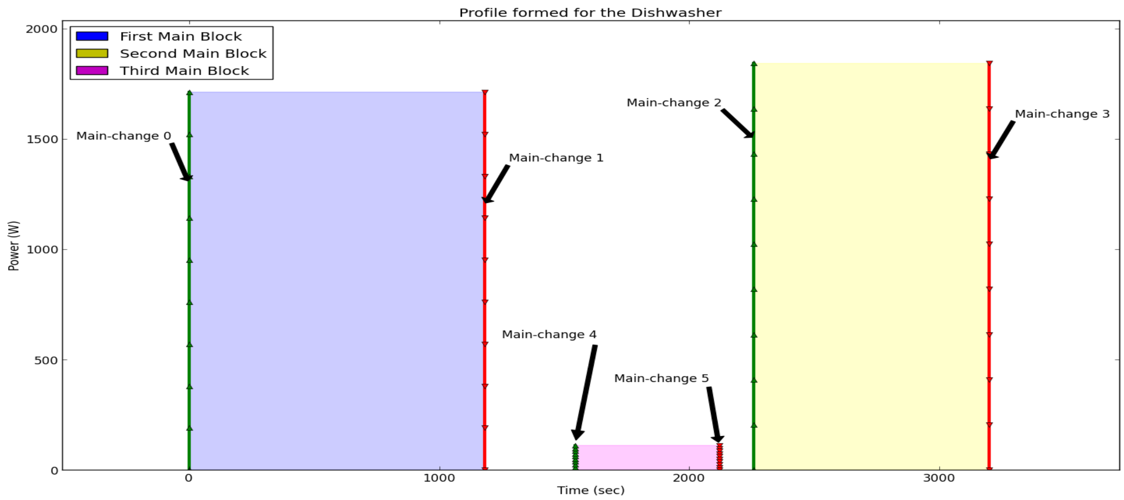


(b)

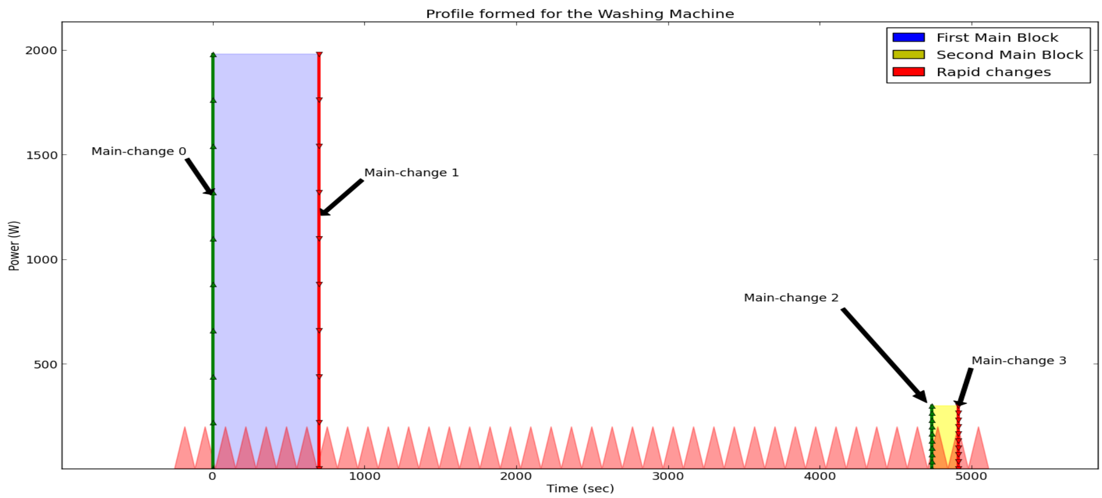


(c)

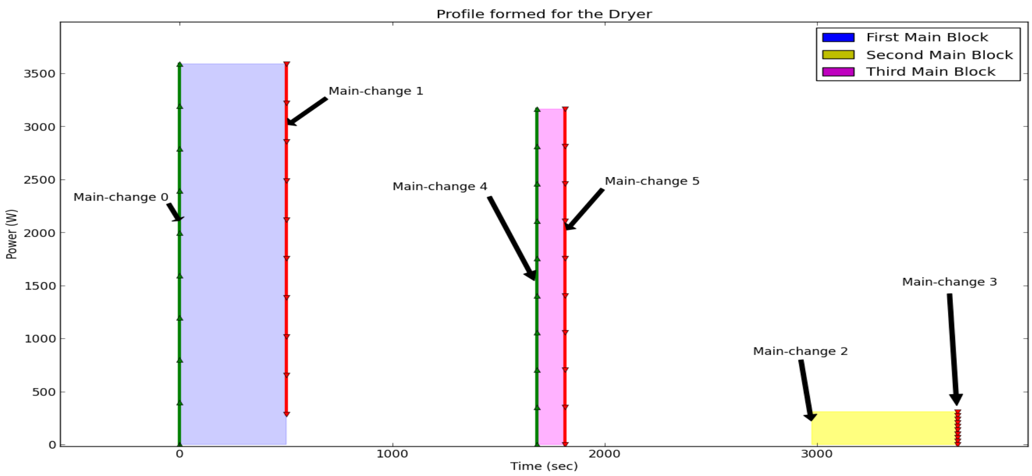
Figure 6.9: Signatures of the three multi-state appliances targeted. (a) Dishwasher, (b) Washing machine, (c) Dryer



(a)



(b)



(c)

Figure 6.10: The main blocks and the main changes of the profile generated for (a) the dishwasher, (b) washing machine, (c) dryer. Up to three main blocks are chosen in order to represent the rough shape of the signatures. The magnitudes and distances between the main changes depicted are the means of the values found while processing the various signatures and are used only to help visualize the blocks' positioning. The actual profile incorporates Gaussian mixture functions to describe the values observed in the signatures.

7 Pattern Recognition

The development of the Profiler module which implements the feature extraction went hand in hand with the development of the Pattern Recognition module, in order to allow for the testing of the technique. The basic functionality of this module was influenced by the design of the prototype system. Naturally, it was extended in various steps in order to utilize the various features extracted by the Profiler.

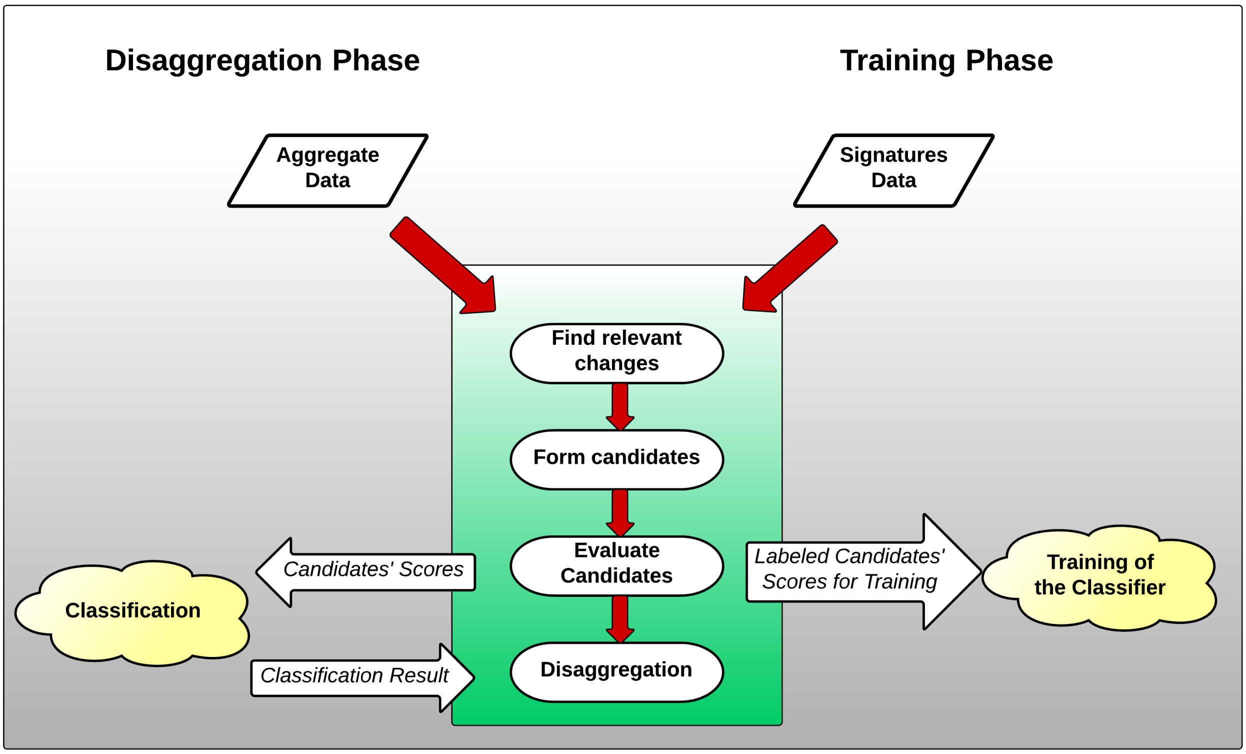
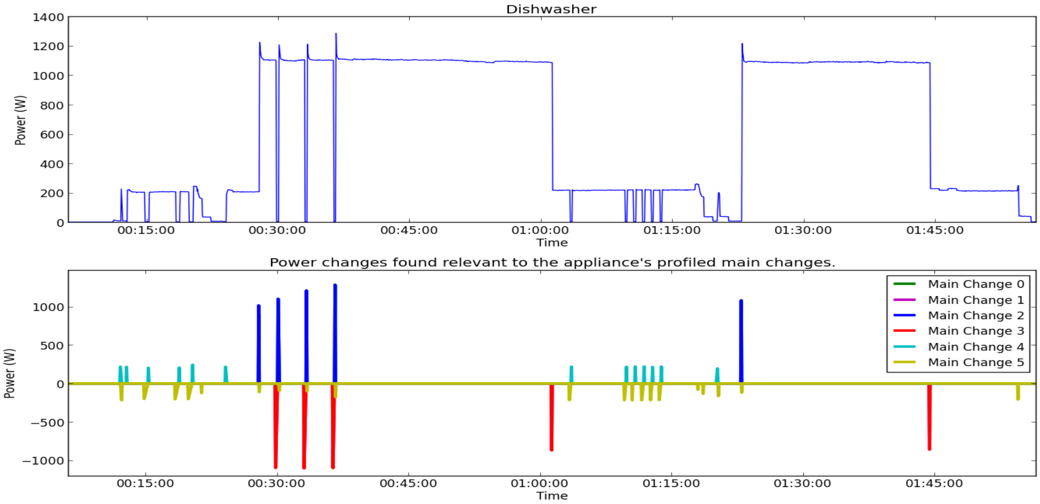


Figure 7.1: Flowchart of the processes performed by the Pattern Recognition module.

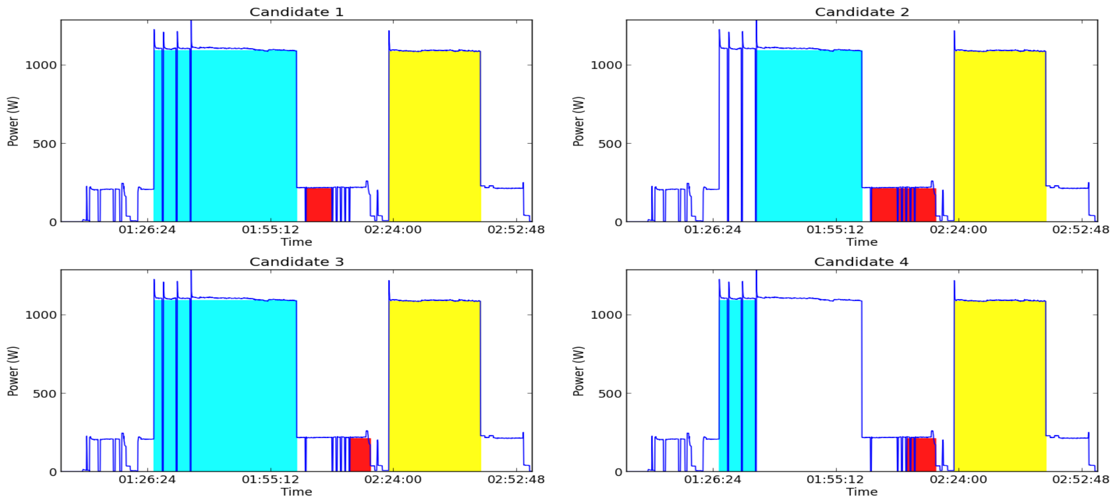
The basic functionality of the module can be described as follows: After the profile of an appliance has been generated, it is provided as input to the Pattern Recognition module. The profile's entries essentially describe to the module how the appliance's signatures look like. By scanning the power signal given as an input, the PR module detects all the power changes that are of relevance to the profile. It then combines the detected changes, creating patterns with similar form to the ones described by the profile. These various combinations are the possible *candidates* for being disaggregated as an operation cycle of the appliance. It then assigns various scores to each candidate, different scores for the various features and characteristics described in the profile. Depending on how close the candidates' features match those of the profile, higher or lower scores are assigned. These scores are then provided to the Classifier.

The Pattern Recognition module plays a double role in the system. Before disaggregation is attempted, the classifier needs to be trained. This is accomplished via the help of the Pattern Recognition. Data with the labelled signatures of appliances are given as input to the PR module, which locates all the candidates in them (multiple candidates can be formed even in one signature, because of the multiple power changes) and assigns scores to them. Candidates with the highest scores are passed to the classifier as positive samples, whereas candidates with low scores are passed as negative samples for training. There is more to be explained about the choice of candidates for training in the relevant section. The second role of the PR module is of course the disaggregation itself, after the classifier had been trained once. After the various candidates are formed in the aggregate signal, they are passed for classification. The PR receives the results from the classifier and finally disaggregates the most confident ones.

7.1 Detection of Relevant Power Changes and Forming the Candidates



(a) The signal is scanned for power changes with magnitudes within the range allowed for the main changes of the profile. This range is described in the profile of the appliance with the minimum and maximum magnitudes that were observed for each main change during the profiling stage. Above are depicted changes detected as relevant to the dishwasher's profile.

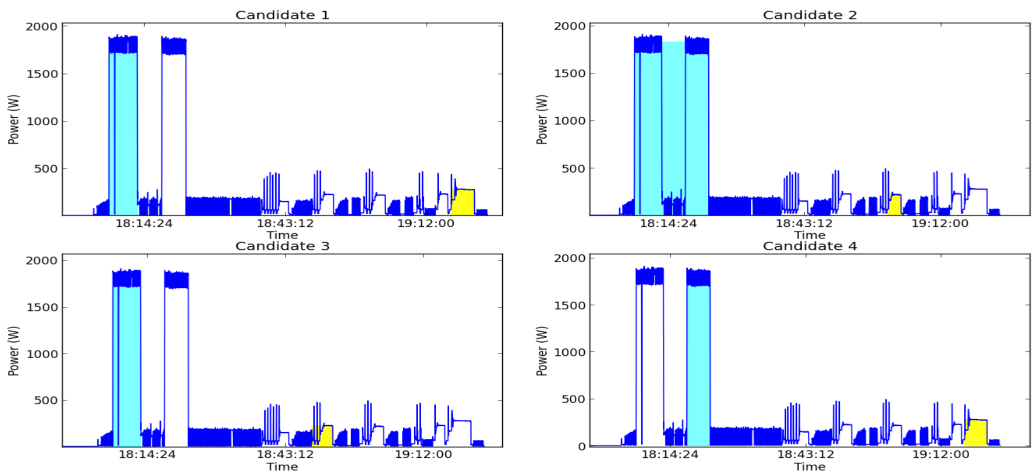


(b) Candidates are formed by the various combinations of the detected power changes. Pairs of positive and negative changes form the main blocks of the candidates. Each candidate is a possible instance of the appliance's operation. After evaluating how close each candidate matches the profile, the closest matches are disaggregated as actual operations of the appliance. Four candidates of the dishwasher are depicted above.

Figure 7.2

The aggregate signal, as commented in section 6.3, is of additive nature. It is created by the addition of the power signals generated by all the operating (and even standby) appliances. When an appliance turns on during the operation of another one, the power consumption adds up and the shape of the aggregate signal is a skewed image, of the two signatures on top of each other. Steady states are for this reason not of much use when it comes to processing the aggregate signal for patterns. What is actually observable in most cases is the power change generated at the change of state of an appliance; thus the disaggregation attempt begins with the detection of power changes in the aggregate, relevant to the profile.

(a) Detected power changes, found relevant to the washing machine's profile.



(b) Four washing machine candidates, formed out of possible combinations of the relevant detected changes.

Figure 7.3

Since, for instance, it is unknown what model of a dishwasher exists in the house, it is needed to find all the changes in the signal that could have been generated by a dishwasher, small or big model. The generic profile generated for an appliance describes the values of the magnitudes that the main changes of the profile take in the signatures of the various models of the appliance. Each main change is described by an individual entry, consisting of not only the density function representing the distribution of the magnitudes as they were extracted from the various signatures, but also of the minimum and maximum magnitudes observed during the profiling process. By inputting the profile to the Pattern Recognition module, it has all the information needed to detect the relevant power changes in the signal. The aggregate signal is scanned for any change with magnitude that falls within the range defined by the minimum and maximum values observed for a main change. The aggregate is scanned both for instant changes and for changes with the form of a spike.

The detected changes are essentially all the possible changes in the signal that could be the main changes of the appliance's profile (see Figure 7.2a). They are subsequently combined in order to form the *candidates* in the signal. The candidates are the patterns in the signal that are possible operations of the appliance. They are initially just combinations of up to six main changes that form the up to three main blocks of the profile. This step can be seen in Figure 7.2b for a dishwasher and in Figure 7.3b for a washing machine. In these figures are presented various candidates that are created even when processing data from a single signature of an appliance. Even in a single signature of a multi-state appliance, it may be the case that several changes appear with magnitudes within the range allowed for the main-changes of the profile. Various candidates may be formed out of their combinations. In an aggregate signal where multiple appliances operate, some of which may consume similar power, the number of candidates becomes quite large. The characteristics and features of all these candidates are compared with the profile of the appliance, in order to find candidates whose characteristics are a close match to the profile.

A last thing to note is the creation of candidates that miss one or more of the main changes of the profile. When the various candidates are created out of the combinations of the detected power changes, additional candidates are created by the module, which miss one or more of the main changes. This enables the module to try to detect an appliance, even if a certain power change is not detected in the aggregate signal (for instance because it is covered by a simultaneous power change which was created by another appliance that turned ON). It also allows for the detection of models of appliances that does not present a certain or more power changes during their operation.

7.2 Evaluating the Candidates - Assigning Scores

The candidates formed initially are just combinations of changes found in the signal, forming the main blocks of the profile. The next step is to compare their characteristics with those described by the appliance's profile, in order to find the candidates that are matching it the most.

Each candidate is assigned a list of scores. Each feature described by the profile is evaluated and a candidate is given a score in the different areas. The list of scores consists of five scores for each main change and three additional scores for the features of rapid changes, repeated parts and the ratio between durations of certain states of the signature. Those various scores are values from zero to one, with higher values indicating that a candidate's feature matches the profile better. This list of scores from each candidate is a sample passed to the classifier for the final decision, whether the particular candidate is a close enough match to be considered a true instance of the appliance.

The overall score of a candidate is calculated as presented below out of the various scores assigned to the candidate in the different areas. It is an indication of how close the shape of a candidate is to the shape described by the profile. The overall score of a candidate is used during the pattern recognition process in order to separate probable and improbable candidates. However it is not the one used for the final classification, but the whole list of scores of the candidate.

$$OverallScore = (weight_0 * Score_{Change-0} + ... + weight_5 * Score_{Change-5} + Scores_{OtherFeatures})/9$$

where:

$$\begin{aligned} Score_{Change-i} = (&Score_{Magnitude-i} + Score_{Correlation-i} + \\ &+ Score_{Distance-OppositeChange-i} + Score_{Distance-OtherChanges-i} + Score_{Type-i})/5 \end{aligned}$$

$$Scores_{OtherFeatures} = Score_{RepeatingParts} + Score_{RapidChanges} + Score_{RatioDurations}$$

The weight of a change is a percentage associated with each main change in the appliance's profile. This percentage expresses the consistency of the particular main change in the signatures processed during the profiling phase; in other words, in what percent of the signatures the particular main change actually existed (also see section 6.4.1). This weight is used in order to lower the influence of the inconsistent changes' scores in the overall score of the candidate.

It needs to be noted here that if an appliance's profile has been formed after choosing less than three main blocks, the number of main changes is smaller, so the list of scores is shorter. Such is the case of the washing machine, where its profile incorporates only four main changes (two main blocks).

7.2.1 Scores for each Main Change

Each candidate consists initially by a combination of up to six power changes detected in the signal. The main part of the candidates' evaluation is to compare these power changes with the characteristics described by the profile for the six main-changes of the appliance's profile. Each of the candidate's changes is assigned separate scores in different areas.

The change's magnitude
When first detecting the relevant power changes in the signal, all that was taken into account is for their magnitude to be between the minimum and maximum magnitude allowed by the profile for the particular main change. This is obviously not enough. In order to evaluate how probable it is for a power change of such magnitude to have been created by a model of the disaggregated appliance-type, the Gaussian Mixture density function of the corresponding main-change entry of the profile is used. As described in 6.4, this density function was formed by fitting a Gaussian Mixture Model to the distribution of the magnitudes observed in the profiled signatures. The probability returned by the density function for the value of the magnitude of the candidate's power change is a factor of confidence that the particular change could have been created by the appliance. If during the profiling phase many models were processed that created main changes of such magnitude, the higher the probability returned for the candidate's change. This probability, after normalization, is the first score assigned to the power change for its magnitude.

Correlation of magnitude with other changes
The detected magnitudes have a random magnitude within the range allowed by the profile. By comparing the changes against the entry of the profile expressing the correlations between the main changes, higher scores are assigned to the changes of candidates that fulfil certain conditions. For instance, a requirement from the profile might be that main-change0 and main-change1 should have about the same magnitude. In another case, it could be allowed that the magnitude of main-change0 should be approximately equal to the some of two other main-changes. Each power change of a candidate is assigned a score depending on whether it fulfils such requirements.

Distances between the candidate's power changes
In a similar way as candidates' changes were assigned scores for their magnitudes, the distances between the various changes of the candidates are evaluated. For the case of the distances, two scores are assigned to each change of the candidate. The first score depends only on the distance between the particular change and the corresponding opposite change that together form a main block of the candidate. This distance describes the duration of a block of the candidate, one of its main states of operation. The score is the normalized probability taken from the density function in the profile that corresponds to the distance between the particular main changes.

A second score is assigned by evaluating the distance of the particular change from all others. In this case the score is the average of the normalized probabilities, which are obtained from the various density functions corresponding to the distances between the couples of changes. This score describes the positioning of the change in relation to the others of the candidate.

Two different scores are assigned, in order to give an increased weight to the distance between corresponding positive and negative changes, that together form a main block of the signature.

Type of each change - Spikes

The change is assigned a score depending on its type, whether it is a normal shaped power change or a spike. The profile contains a percentage that expresses in how many signatures a main change had been observed to be a spike, out of all the signatures processed during the profiling phase. This percentage is the score assigned to the candidate's change if it is indeed a spike, or its complement is assigned if it is of a regular shape.

7.2.2 Scores for Other Features

Scores for the existence of Repeating Parts

Three further features are evaluated for each candidate, the first of which is the existence of repeating parts of the signature. As explained in section 6.7, during the operation of certain appliances it is observed that some of their elements, for instance the heating element of the dryer, turn on and off several times. These parts of the signature are modelled as repeats of a main block. If such behaviour has been observed for an appliance, its profile contains an entry with the number of the main block that is repeated, as well as the area that the repeats occur. This area is described with the use of distances from the main changes of the profile. Finally, the entry describes the type of the repeats. As explained also in section A, the type describes the number of the repeats usually observed in the various signatures of the appliance, as well as whether they have fairly similar durations and time intervals between them.

In order to evaluate this feature for every candidate, using the distances provided in the profile and the position of the candidate's power changes, the area that they are supposed to occur is defined in the signal. Any changes that are detected in the specific area of the signal with magnitudes similar to the candidate's change specified by the profile are extracted. These changes are then grouped together to form blocks, the repeating parts. If their number and type is according to the type described by the profile, a score of 1.0 is assigned to the candidate for this feature and the repeating blocks are added to the candidate as parts of it.

Existence of rapid changes

If during the profiling phase it has been found that there are periods of rapid changes in the appliance's operation, an entry of the profile describes the areas that the rapid changes begin and end, in relation to the main changes. If there is such a requirement from the profile, the Pattern Recognition module locates the areas described by the profile in relation to the changes of each candidate. It checks if there are indeed rapid changes beginning and ending there. If this is the case, a score of 1.0 is assigned to the candidate for this feature, or a zero if the rapids are not found.

Ratios between durations

A last possible feature of an appliance's profile is the ratio between the durations of certain states of the appliance's signature. If there is such a feature, the relevant entry of the profile contains the characteristics of the GM density function. The ratio between the distances of the candidate's changes that correspond to these durations is calculated. The normalized probability obtained by the density function for the candidate's ratio value is the score assigned for this feature.

7.3 Candidates for Training of the Classifier

The Pattern Recognition module plays an important role in the training of the classifier. For this phase, the data with the signatures for training are passed as input to the Pattern Recognition module, which processes each signature individually. The relevant changes are detected in each signature and then it proceeds to form candidates out of their combinations. As discussed in 7.1, even when processing a single signature of the appliance there are various candidates being formed. In order to accomplish the positive training of the classifier, the PR module follows the steps presented above and assigns scores to all the candidates. It then calculates the *overall score of a candidate*, as introduced in 7.2.

The overall score of a candidate is an approximation of how close the characteristics of the candidate match those of the profile. In order to extract the positive training samples for the classifier, out of every signature of the appliance the candidates with the highest overall score are chosen. The set of scores (and not the overall score) of each candidate is a training sample for the classifier. At this point it is interesting to notice a convenience of the use of candidates and scores. As mentioned before, even in a single signature there are more than one candidates created. This offers on option for the positive training: The classifier can be positively trained on either the top ranked candidate from each signature or , alternatively, in more than one high ranked candidates, as long as they have an overall score close enough to the highest one. The second option seems rather unusual. However, it is rather convenient as it allows for the 'artificial' creation of positive training samples with slight differentiation. This allows the classifier to generalize more efficiently than when trained on only one positive sample from each signature, because of the fairly limited available data.

The negative training can be accomplished in a similar way as the positive. In this case the use of candidates also allows for a rather unusual choice. The first and obvious choice is to input data with signatures of other appliances to the PR module, extract the candidates and provide their scores to the classifier with negative labels. However, since the characteristics of those candidates are usually very far from the profile's characteristics, the scores are really low. This attempt did not give good enough results, as the classifier does not manage to adjust its decision boundary appropriately. The alternative is to extract candidates from the signatures of the actual (true) appliance that the classifier is being trained to classify. In this case, the PR module gives as negative training samples to the classifier the scores from candidates that are further below the top ranking ones, but still have high scores in certain characteristics of the profile. This approach led to better results, although it was still not the final one adopted. The final choice of classifier does not actually require negative training. This will be further explained in the section about the Classifier module.

7.4 Disaggregation

The final step is the actual disaggregation. Out of the various candidates, some are classified by the Classifier module as true operations of an appliance, others as false, depending on how close their scores are to the trained positive ones. There is however one final thing to take into account. As explained in section 7.1, the amount of candidates created by all the combinations of the relevant detected power changes in the signal is big, with some of them being very similar with each other. For instance some of them have common five out of six main changes and have almost all features in common. In some cases this results to a positive classification of more than one candidates in the same area of the data.

In order to solve this problem, the disaggregation is achieved similar to the one described for the prototype version. Inspired by the 'sliding window' technique used in OCR, the data are swiped from the beginning to the end. If a positively classified candidate is met, it is checked whether within its duration there is also some other positively classified candidate. If it is not the case, the lone positive candidate is disaggregated as a true operation of the appliance, its position and energy consumed is reported for evaluation. The window continues swiping the data from the point that the disaggregated candidate's duration is finished. In the case of more than one positive candidates within the proximity, the classification score returned from the classifier is checked. Only one of candidates is disaggregated, the candidate with the highest reported classification score.

7.5 Computational Efficiency Problems and Optimization

One of the biggest issues faced while designing the pattern recognition technique had to do with computational efficiency. As commented in previous sections as well, the aggregate signal is a rather complicated signal. The operation of the numerous appliances in a house create a wide variety of patterns in the aggregate. In addition, different appliances may consume similar amounts of power. As a result, there are areas in the aggregate signal where the PR module may detect a big number of power changes that are relevant to the profile. A big number of relevant changes can create a vast amount of candidates out of all the possible combinations.
If m main-changes are used in an appliance's profile, the number of candidates formed is:

$$candidates = n_0 * n_1 * ... * n_{m-1} * n_m$$

where n_i is the number of changes detected relevant to the main change i. If we assume n detected changes relevant to each main change, the number of candidates can be expressed as n^m.

The fact that the number of candidates increases exponentially with the number of main changes used in the profile, explains the choice of creating the profile by choosing only up to three main-blocks, as explained in the previous section. With six corresponding main changes, there are still too many candidates being formed occasionally in the signal. Evaluating all the candidates is a time consuming process. This bottleneck created various problems throughout this work and various attempts were made in order to achieve some better efficiency.

7.5.1 More Efficient Candidate Creation

The initial attempt to reduce the number of candidates was to take into account the distance between the detected changes in the aggregate signal. The only candidates that are formed should be created by combinations of power changes that do not have a smaller or greater distance than what is allowed by the profile. For instance if during the profiling phase of a dishwasher, the shortest heating cycle observed for the various models of dishwasher lasted X minutes and the longest one Y minutes, no candidates should be formed where the corresponding power changes in the aggregate have a distance outside that range.

The profile of the appliance contains entries for the distances between each couple of main changes. These entries contain the minimum and maximum value that each certain distance has been observed to take among all signatures of the appliance profiled. Using this information, when processing the various combinations of power changes, candidates are only formed by combinations where the distances of the power changes fall within the corresponding limits. This initial requirement limited to a great extend the number of candidates. Still, it was not enough in certain cases, where the operation of various appliances at the same time would create a big amount of power changes.

7.5.2 Removal of Malshaped Candidates

The next attempt to reduce the number of candidates was to eliminate in early stages of the process those that formed malshaped patterns. Such a candidate is presented in Figure 7.5. The candidates' power changes form the various main blocks of the appliance's profile, which essentially represent the operation of certain elements of the appliance. Since the height of the block is the power consumption of the appliance during that shape, the power level does not make sense to drop below that level for the duration of the block. When the candidates are formed out of all the combinations of the power changes, various such mal shaped candidates are formed. By detecting whether the power drops below acceptable levels during the duration of a block, such candidates can be eliminated in early stages of the process, gaining a lot of computational time in the next

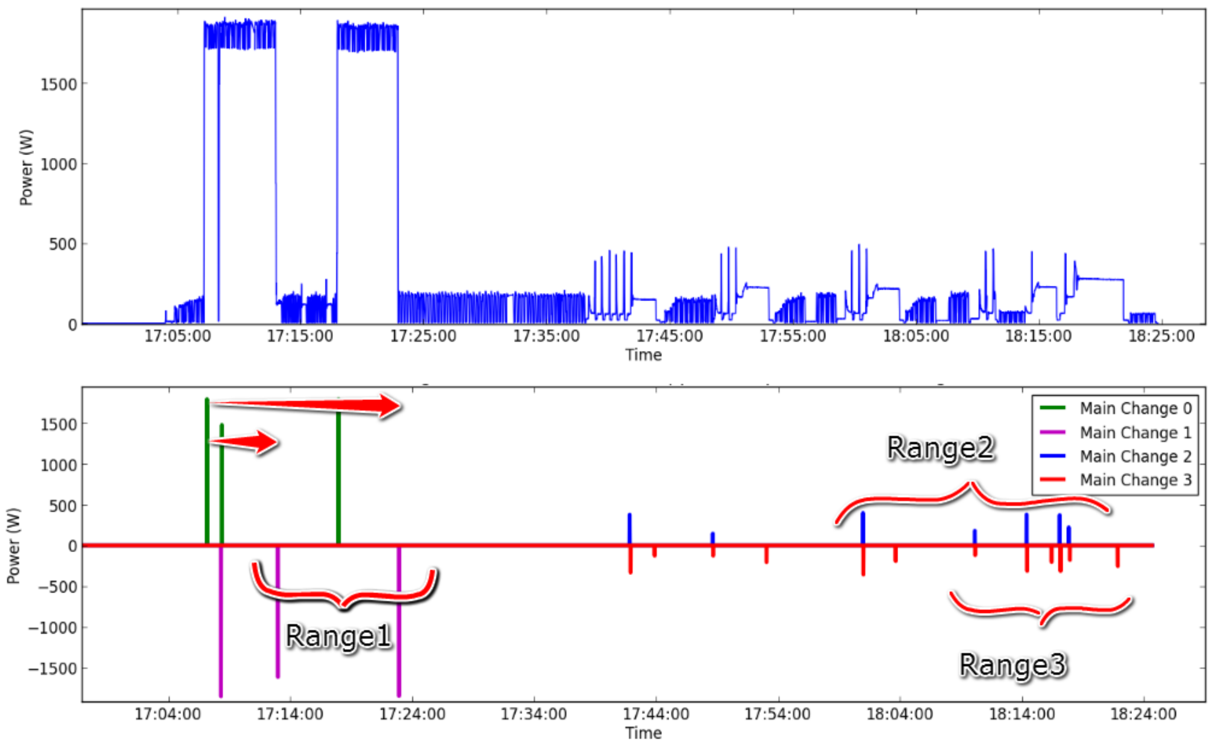


Figure 7.4: Candidates are only formed by combinations of power changes the distances of which are within allowed range by the profile. The above picture presents an example of allowed ranges for the various main changes. A candidate will be created by combination of changes that fall only within those ranges. For instance, the first detected positive change (which corresponds to main-change0 of the profile) will only be combined with negative changes within range1 to form the first block of candidates.

evaluation stages [1].

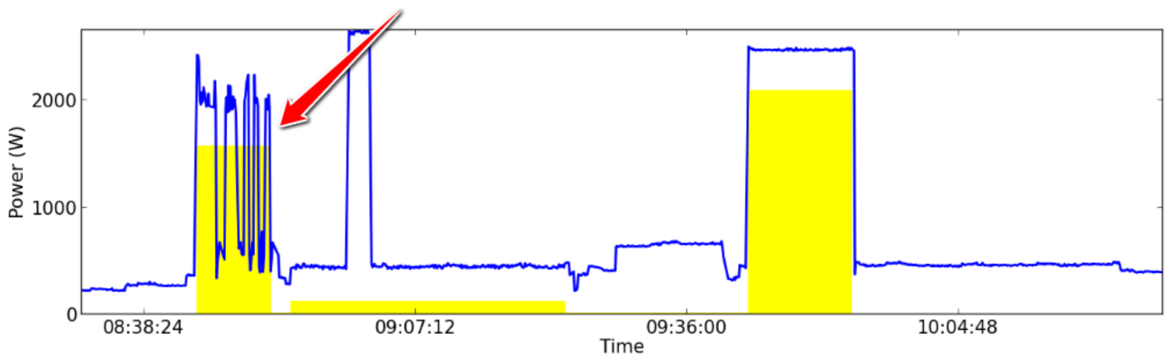


Figure 7.5: In order to reduce the number of candidates, the ones with mal shaped blocks are eliminated in the early stages of the process. Above are presented the three blocks of such a candidate during the disaggregation of a dishwasher. The first block of the candidate is problematic, with intermediate power changes during its duration that drop the power level below acceptable.

[1] During the profiling process, short breaks in an appliance's state were overlooked, in order to extract the main shape of the signature (see section 6.1). The same threshold is applied in this case, disregarding very short intermediate breaks in a block

7.5.3 Clearing Low Score Candidates

In a further attempt to reduce the number of candidates and achieve shorter processing times, candidates' overall score is checked halfway through the evaluation process. After the assignment of scores to each change of a candidate and before the evaluation of the rest of the features, their current overall score is checked in order to remove low scoring candidates. By removing the candidates with overall score less than the 50% of the highest score in the scanned area of the signal, the number of candidates is reduced by approximately 55% in most cases, shortening significantly the rest of the process. The threshold of 50% has not been found to cause problems by eliminating wanted candidates, since it is unlikely for an appliance's model to score that low in all areas evaluated for the main changes. An increased threshold however increases this risk significantly, reducing the ability of the technique to generalize to unknown models.

7.6 Limitations Introduced

The design of the pattern recognition technique introduced some limitations to the system. The first limitation influences the detection of the relevant changes in the signal. In order for the power changes in the signal to be considered as relevant to the profile's main changes and subsequently be taken into account for the creation of the various candidates, their magnitudes have to be between the minimum and maximum magnitudes found for a main change when creating the profile of the appliance. If, for instance, a dishwasher model is smaller than all of the models processed during the profiling, the power changes it creates will not be detected, as they will have a smaller magnitude than the allowed by the profile. The disaggregation attempt is bound to fail from the very first step.

The second limitation is of similar nature. In the second step of the process, in order to reduce the number of candidates created, the distances between the changes is taken into account. Candidates are created only from combinations of power changes that their distances fall within certain limits, provided by the appliance's profile (see section 7.5.1). Just like the case explained for the magnitudes, if a model of an appliance has a shorter duration of operation than all the models processed during the profiling process, proper candidates will not be formed and the disaggregation will fail.

This problem is alleviated by adjusting the limits of the magnitudes and distances limits by a small percent, for instance by 20%, allowing the system to be more flexible when detecting unknown models. Still, disaggregation of an even smaller or larger model of the appliance would fail. This problem is not a limitation of this approach alone. In every pattern recognition approach, classification of a sample depends a lot on a complete enough training set. A sample that is far different from the ones observed during training is very probable to be misclassified.

Finally, an important limitation has been introduced to the system because of the design of the disaggregation phase. The system cannot disaggregate two appliances of the same type if both operate simultaneously. For instance the module cannot disaggregate multiple dishwashers in a laundry room, operating at the same time.

8 Classification

The classifier is the module which completes the system. After the aggregate signal has been processed by the Pattern Recognition module and the various candidates have been formed and evaluated, the classifier is called to make the final decision whether a candidate represents a true operation of the disaggregated appliance or not. This is achieved by comparing the scores that each candidate received in the various aspects of the profile, against the trained samples, scores acquired from the signatures of various models of the appliance during the training phase. If the candidate's scores for the various features are close enough to the ones trained from the signatures, the candidate is positively classified. The results from the classifier, a positive or negative label for each candidate, are returned to the PR module in order to do the final disaggregation. The choice of the classifier, its training and functionality are discussed below.

8.1 Choice of the Classifier

During the early stages of the system's development, a naive version of a classifier was used as a place-holder in order to allow for debugging. This simplistic classifier was simply calculating the overall scores of the candidates, and classified as true operation of the appliance the highest ranking candidate in the region of the data, as long as its overall score was over 50%. This version proved really useful for debugging and was actually still used for that reason even during the final stages of the work, since it is guaranteed to return the candidate with the highest score without any training. Of course it fails to eliminate any high scoring false positives, problem alleviated after the adoption of an actual classifier.

The investigation for a classifier began by trying linear classifiers such as Logistic Regression with first order features, as well as the linear Perceptron. These linear classifiers failed to provide any good results. This seemed rather unexpected at first, since the scores provided to the classifier are defining a close match when they take high values, negative classification when they are low. One would expect that by simply adjusting the weights of a linear decision function would suffice. However when the investigation for the classifier choice was performed, the system was still not in its final form, with various problems in the scoring process. This did not allow a linear classifier to perform well.

More classifiers where tried in the following attempts, including a Logistic Regression classifier with the use of second and third degree features, as well as Support Vector Machines classifiers, with the use of various kernels. All these attempts gave fairly positive results after some configuration. The Logistic Regression classifier was still used in the final stages of this work for various testings. The final choice however was the use of a One-Class-SVM by Scholkopf, implemented in the libsvm library.

One-Class Support Vector Machine

The One-Class-SVM classifier is a one-vs-all classifier (see section 2.5) with a very interesting and convenient attribute. The particular type of SVM does not need negative training. It is only trained on positive samples and adjusts a hyperplane in the multi-dimensional feature space in order to envelop those positive samples. New, unknown samples are classified as positive or negative, depending in which area of the feature space they fall in; positive if they fall in the inner side of the hyperplane, negative if on the outside. At the risk of oversimplifying, the One-Class-SVM 'learns' the values that the features of the positive training samples take. The area in the multidimensional feature space that the positive candidates belong is defined by the coordinates given by those values. It then classifies a new sample whether or not its features are close to the trained ones. This classifier is used for novelty detection and anomaly detection problems, where the positive state is well known. Everything that does not match the known characteristics well enough is an anomaly and classified negatively. The nature of this classifier matched particularly well the nature of our problem.

The One-Class SVM can be used with various types of kernels. Various of them were tried during a short validation phase. The kernels that gave the best results were the RBF (Gaussian) kernel and the third degree polynomial kernel. Both types of kernels performed well in classifying the anomalies in the scores, giving very few false positives even from the very first tests. The RBF kernel proved to be a little more strict than the polynomial, misclassifying as false approximately 20% of true candidates more than the polynomial. The polynomial kernel seemed to be able to generalize better, so it was the final choice, used for the final evaluation of the approach.

Disadvantage of the use of Support Vector Machines

Classifiers such as Logistic Regression and Neural Networks are able to report a probability score, describing their confidence. Their confidence score is the output of their sigmoid function. Unfortunately the Support Vector Machines do not have a similar property. SVMs do report a classification score, which is the distance of a sample from the hyperplane separating the multidimensional feature space in the areas of positive classification and negative classification. A positive high score means positive classification with high confidence. Unfortunately this score is non-normalised, with a range that depends on the decision function, subsequently on the training set and the parameters of the SVM. A method has been developed in [45] for estimating such a probability. Because of the fact that such a probability would not be of much use to the system with the current design and also to avoid making the process of the candidates even more computational expensive, the technique was not used.

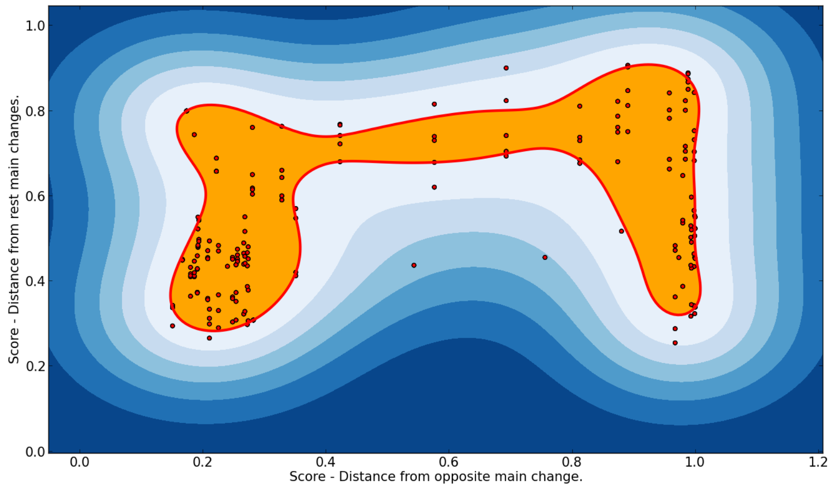
8.2 Training and Classification

The One-Class-SVM is a one versus all classifier. It is trained to classify positively one and only class. Just like one profile needs to be generated for each appliance type, a separate classifier object needs to be trained for each type of appliance. This process needs to be done only once for each appliance. Each training sample for the classifier is the list of scores that a candidate has received during its evaluation by the Pattern Recognition module. The training set consists of such scores from various candidates, the top ranking candidates extracted from the signatures of the appliance. The process of extracting the appropriate candidates for the training set is described in section 7.3.

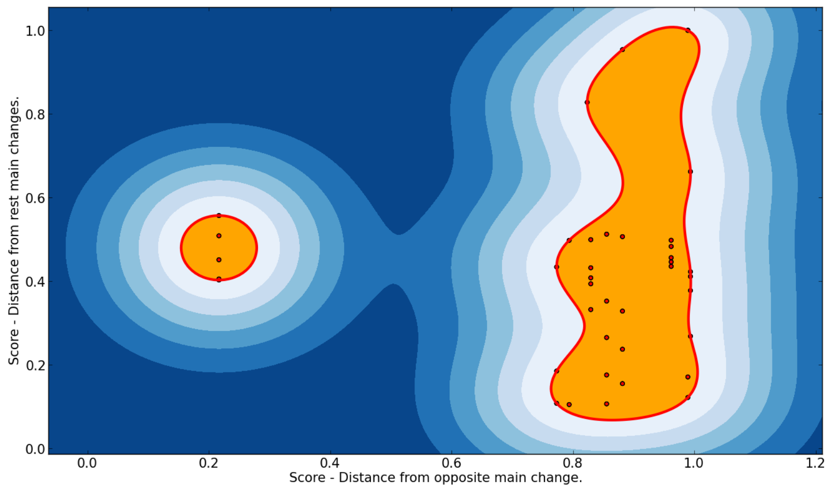
Depending on the number of main changes in the profile, the number of features that the classifier gets trained on (which is the number of scores for each candidate) varies. For each main change of the profile five scores are assigned, with additional scores assigned for the existence of rapid changes, repeated parts in the signature and the ratio between the durations of certain states in the signature with similar power values (see section 7.2). This fairly big number of features, up to thirty-three features for a six-main changes profile (three main blocks chosen), comprises the feature space of the classifier. Within this multidimensional space the classifier is called to adjust the hyperplane decision boundary.

Visualising a feature space of that many dimensions is impossible. It is possible however to get a feeling of the task that the classifier performs by visualising pairs of the features in the two-dimensional space. In figure 8.1 are presented the decision boundaries of an One-Class SVM with an RBF kernel, trained on a dishwasher (a) and a dryer (b). The classifier separates the feature space into two areas. Candidates that belong in the inner side of the decision boundary, closer to the trained samples, are positively classified. Candidates with scores that put the in the outer side of the decision boundary are detected as anomalies and are negatively classified.

In the same figure can also be observed an interesting point about the use of a big number of features in the profile. Some of the features seem to not help particularly in the distinction between a positive and a negative candidate. For instance, in the case of the dishwasher the score for the distance of the first main change from the opposite negative change takes a big range of values, with the classification area of the SVM extending from one side of the feature space to the other. The same feature,however, does make a difference in the case of a dryer. The SVM 'learns' during the training that the various models of dryers present either a high or a low score in this feature. This can be compared with the density functions for the particular feature, extracted during the profiling process, presented in figures B.6(a) and B.8(a) (As a explained in section 7.2.1, a score for a feature is the normalised value acquired from the density function generating by fitting a GMM to the values extracted during the profiling process).



(a)



(b)

Figure 8.1: The decision boundary of an One-Class SVM RBF classifier (red), in the two dimensional space of the features: score of the first main change for the distance from the opposite main change (x-axis), score of the first main change for its distance from the rest of the main changes (y-axis). Candidates that fall inside the decision boundary are positively classified. An important thing to notice is that some features might not make much difference for one appliance, but they may do for another. For instance, the first feature does not help reduce the length of the positive classification area for the case of the dishwasher (a). However, in the case of the dryer (b) it does lead to such reduction.

9 Evaluation

The designed technique and system were tested separately for the three targeted multi-state appliances, the dishwasher, the washing machine and the dryer. The aim was to investigate the generalization capabilities of the technique by attempting to disaggregate a model of each appliance that the profiler and the classifier were not trained on. The evaluation was performed on two datasets: The REDD dataset, sampled every 1 second. The second dataset tried is the one collected by Jack Kelly, co-supervisor of this research. This dataset was available in two granularities, sampled every 1 second and every 6 seconds. For the cases that Jack Kelly's dataset was used (dishwasher and washing machine), both the sampling rates were tried in order to test whether the system is influenced by the sampling rate. The detected events were very similar in the two cases, so they are presented only once below.

9.1 Metrics

As discussed in section 2.2.4, presenting the overall efficiency of a disaggregation technique is not a trivial task. The false positives detected by a disaggregation system (patterns that are mistakenly detected as operations of the appliance) make it difficult to evaluate its overall efficiency. For instance, a system might detect all the operations of an appliance in the aggregate signal but at the same time falsely report many other patterns as being operations of the appliance. In order to provide a complete and coherent view on the system's efficiency, the following metrics are used to describe the performance on a dataset:

- Disaggregation Accuracy:

  $$\frac{\text{True Positives}}{\text{Actual Events}}$$

  This metric (introduced in [22]) describes the percent of the actual operations of an appliance that the system managed to detect.

- Detection Accuracy:

  $$\frac{\text{True Positives}}{\text{True Positives + False Positives}}$$

  This metric describes the efficiency of the system taking into account the effect of false positives detected (also introduced in [22]).

- Percent of Energy Consumption Correctly Disaggregated:

  $$\frac{\text{Total Energy Consumption of True Positives}}{\text{Total Energy Consumption of all Actual Events}}$$

  The ultimate goal of a disaggregation attempt is not only to detect when an appliance operates, but also calculate the energy it consumes. This metric reports the percent of the total energy consumed by the actual event which is accurately disaggregated (not taking into account false positives).

- Percent of Energy disaggregated from Detected Events:

  $$\frac{\text{Total Energy Consumption of True Positives}}{\text{Total Energy Consumption of Corresponding Actual Events}}$$

  This last metric is used in order to describe what part of the energy consumption of the detected events the system manages to disaggregate. It allows to estimate the part of the operation that the system leaves out, due to undetected parts of the signature.

9.2 Evaluation for the Dishwasher

Disaggregating the dishwasher was attempted on two aggregate signals. The first belongs to a house from the dataset collected by co-supervisor of the project, Jack Kelly. The second signal is from House-2 of the REDD dataset, in order to allow for comparisons with future work.

Before the disaggregation attempts, the profile of the appliance had to be formed, as well as train the classifier. In both cases, the profile was created by all the models of dishwashers in the available datasets, except the particular model that was to be disaggregated. Because not all models have the same number of signatures available, only up to four signatures were used from each model, in order to have approximately the same influence from each model to the profile and training sample. Same thing applies for the training of the classifier. This is an important point, in order to prove the capabilities of the technique in generalizing for unknown models. Please note that this is not discussed in lots of cases of previous related work, with the possibility that the designed techniques and the reported results may come from training and testing a system on the same dataset.

The system was tested at first on measurements from 25 days from the first house. For the disaggregation attempt on the house from the REDD dataset, the aggregate signal was scanned for 18 days in total. The results from the two attempts are presented in table 4.

Table 4: Dishwasher disaggregation: Results

	Jack Kelly's dataset: House-1	REDD: House-2
Actual Number of Signatures	9	6
Actual Power Consumption	11.55 kWh	4.05 kWh
Number of Signatures Disaggregated	10	6
Number of True Positives	9	6
Number of False Positives	1	0
Energy Consumption Correctly Disaggregated	10.16 kWh	3.6 kWh
Disaggregation Accuracy	100%	100%
Detection Accuracy	90%	100%
Percent of Energy Correctly Disaggregated	87.9%	88.9%
Percent of Energy Disaggregated from Detected Events	87.9%	88.9%

The results from both houses are satisfying for the dishwasher. The system manages to detect all the instances of the appliance's operation, even though neither the profile was formed, nor the classifier was trained using the models being disaggregated. The technique seems to successfully generalize to other models for the case of the dishwasher. A very important point to notice is the very low number of false positives. With only one false positive found in one of the houses, the reliability of the technique seems rather high. This is the result of using a big set of features in order to detect the pattern of the disaggregated appliance. Although many candidates-patterns are found in the aggregate signal, the big number of features lowers the probability that a random pattern will match an actual dishwasher's operation in most of them. As a result, the classifier detects those missing features of the candidate-patterns as an 'anomaly' in comparison to the trained samples, resulting in the low number of false positives.

Inefficiencies

Very interesting are the results for the amount of energy disaggregated and the reasons behind this percentage. Figure 9.1 presents a disaggregated instance of a dishwasher. It is easy to notice

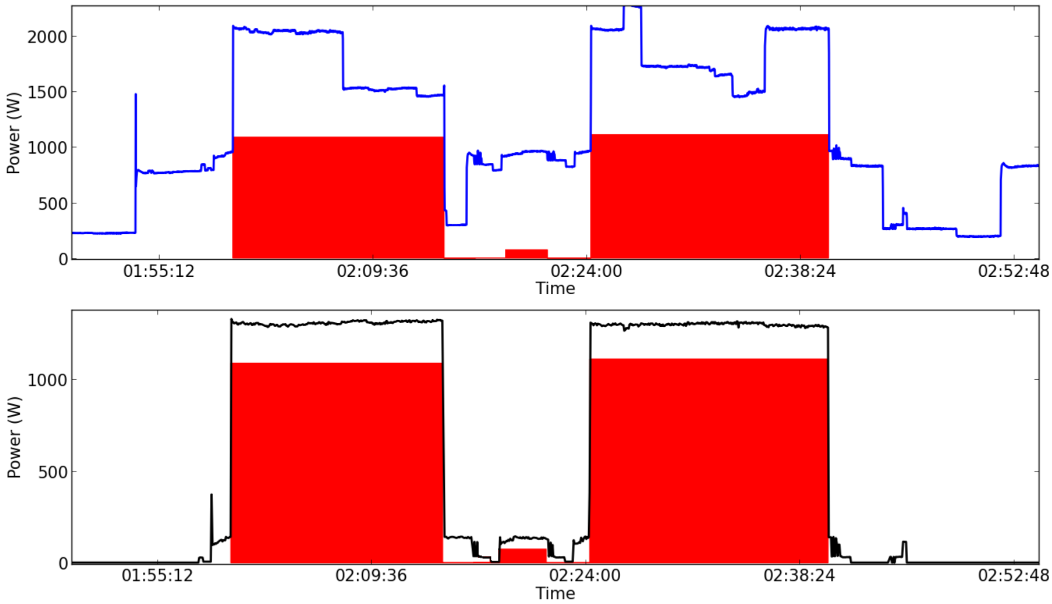


Figure 9.1: The aggregate signal (top), the signature of a dishwasher (bottom) and the disaggregation's result (red). The pattern recognition fails to detect some parts of the operation. The reason is that the low-power parts on either side of the heating cycles are not among the three main blocks chosen to represent the rough shape of the appliance's signature. They are not represented as repeats of the third main-block (low-power middle block) of the signature either, because the Profiler failed to detect them consistently in the dishwasher's signatures. This led to inability to represent them in the dishwasher's profile and subsequently to detect them.

the parts of the signature that are not disaggregated, the small blocks on the left and right of each heating element's block, as well as a part of the blocks that correspond to the two heating cycles. It is easy to find the explanation for this failure if we look back at how the Profiler describes the rough shape of an appliance's signature. A visualisation of the profile's chosen main blocks have been presented in Figure 6.10 (a). The Profiler detects up to three consistent blocks of the appliance's signatures. The low-power parts not detected during disaggregation are not among the three main blocks of the appliance's profile. These parts could be represented as 'repeated parts' (see section 6.7), repeats of the third main block of the profile (the low-power block in the middle part of the signature). However, during the creation of the profile, the blocks in the first and last part of the signature were not found in the majority of the signatures of the models. For that reason they were not added to the profile as repeated parts either, leading to inability to detect them. The missing top part of each heating cycle is a consequence of not detecting the low-power blocks that precede them. During the pattern recognition, the power changes that are detected are the ones leading from the lower power state to the high power state. Since the preceding low-power parts of the operation are not detected, the heights of the two main blocks corresponding to the heating cycles are lower, leading to reduced disaggregated energy.

9.3 Evaluation for the Washing Machine

The process followed for the evaluation on the washing machine is similar to the one followed for the dishwasher. The washing machine's profile was formed out of the signatures of all the models available in the datasets, but the signatures of the model that was to be disaggregated. The One-Class-SVM classifier was trained on the same signatures. Just like the case of the dishwasher, up to four signatures were used from each model of washing machine, in order for the profile formed and the classifier's training to be equally influenced by each model.

As explained in section 3.2.1, the washers in the REDD dataset are of a different type, with a very different operation. These washing machines were not targeted in this work. For this reason the aggregate signals from the REDD dataset could not be used for this part of the evaluation. Instead, the evaluation was performed using the aggregate signals from two different houses in the dataset collected by co-supervisor of this work, Jack Kelly. The system was tested on measurements from the first house of 17 days length. The aggregate signal from the second house was scanned for 22 days. The results from the two attempts are presented in table 5.

Table 5: Washing machine disaggregation: Results

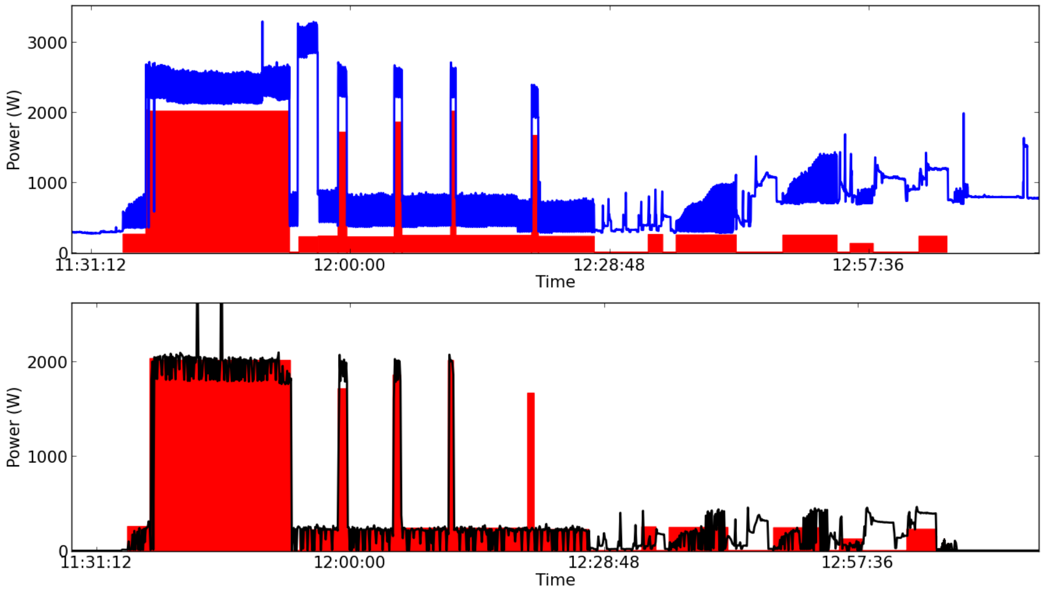
	Jack Kelly's dataset: House-1	Jack Kelly's dataset: House-2
Actual Number of Signatures	10	11
Actual Power Consumption	9.75 kWh	4.93 kWh
Number of Signatures Disaggregated	9	7
Number of True Positives	9	7
Number of False Positives	0	0
Energy Consumption Correctly Disaggregated	7.55 kWh	2.71 kWh
Disaggregation Accuracy	90%	63.6%
Detection Accuracy	100%	100%
Percent of Energy Correctly Disaggregated	77.4%	55.0%
Percent of Energy Disaggregated from Detected Events	86.0%	86.5%

Disaggregating the washing machine from aggregate signal of the first house was performed with high disaggregation accuracy, missing only one of the washer's operations. The system was not as efficient on the second house however, missing 4 out of 11 operations. The reason behind this is that the operation of the second house's model has the shortest duration out of all the models in the datasets. As a result, the classifier has less confidence in identifying its operation, classifying some true events as anomalies. Even though some true events were missed, in both houses the detection accuracy of the system was 100%, avoiding to detect any false positives. Moreover, the percent of energy disaggregated from the detected events is fairly satisfying.

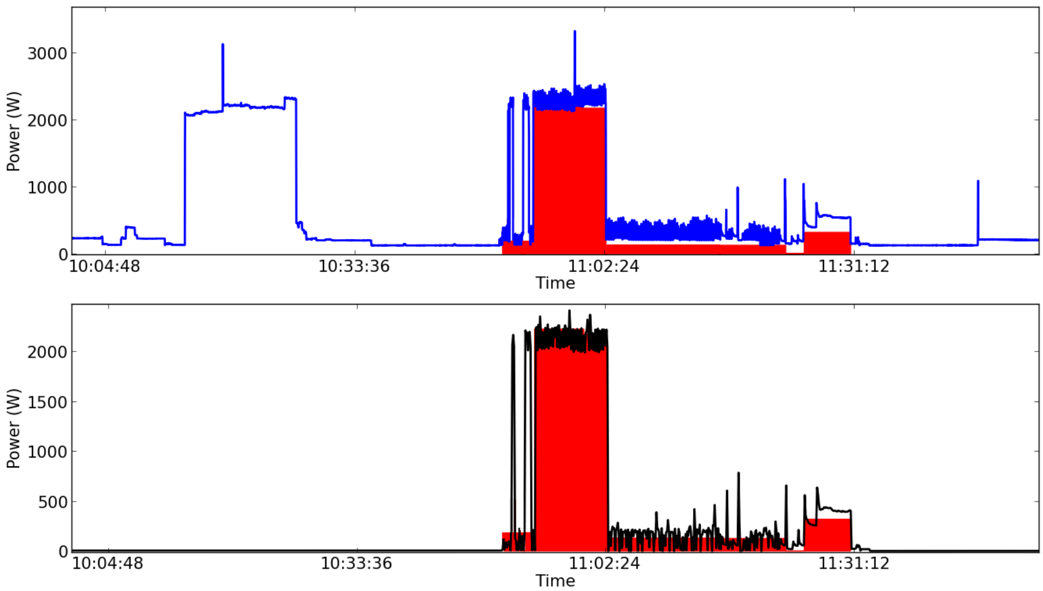
Inefficiencies

Aside the overall percent of disaggregated energy, checking the energy disaggregated for each individual signature that was successfully detected reveals an average percent of 86.0% for house-1 and 86.5% for house-2. Certain parts of the washing machines' operation were consistently missed. In figure 9.2 are presented disaggregated instances from house-1 (a) and house-2 (b). The washing machine of house-1 presents multiple rinsing cycles. Similar with the case of the dishwasher, these parts are not represented in the appliance's profile. These parts of the signature could be represented in the profile as repeats of the final rinsing cycle, but during the evaluation they have not been observed in the majority of the profiled signatures. Additional tests with a different set of signatures were performed, where the majority did exhibit that behaviour and those parts were modelled by a 'repeating parts' entry in the profile. The results were improved, but not by much. The Pattern Recognition module still had problems detecting them. The reason is that the shape of these parts does not really resemble a 'block', the kind of shape that is easily detected by the technique.

The second problem observed concerns the detection of the short activations of the heating element. These short, high-power consuming parts of the signatures are modelled in the profile as repeats of the first main block of the signature, the main heating cycle. Those repeats were mostly observed in the middle part of the signatures during the profiling phase. The washing machine model of house-2 is one of the few models in the datasets that exhibit such behaviour at the beginning of the signature, before the main heating cycle. Since this behaviour is not observed for most models, the profile does not include a 'repeating parts' entry for the area at the beginning of the signature. As a result, those parts are not modelled, they are 'unknown' for the Pattern Recognition module and are not detected. Another problem regarding the 'repeating parts' model of the system is observed in figure 9.2 (a). The three first repeats are successfully detected. However an unknown appliance with similar power consumption and similar duration turned ON within the area that is specified by the profile that these repeats occur. This unknown device is therefore detected as a forth repeat, mistaken for a part of the washing machine's operation.



(a) Aggregate signal (top) and signature of the washing machine (bottom) in house-1, with the disaggregation's result (red).



(b) Aggregate signal (top) and signature of the washing machine (bottom) in house-2, with the disaggregation's result (red).

Figure 9.2: Detected instances of the washing machine's operation in (a) house-1, (b) house-2, with problematic detection of certain parts of the signature. Certain parts of the operation have not been consistently observed in the signatures of the various models. For that reason they have not been added to the profile, preventing their detection as the Pattern Recognition does not 'know' of their existence. Moreover, an unknown device is falsely detected as being part of the washing machine's operation, because of its similarity with true parts of the signature.

9.4 Evaluation for the Dryer

The last targeted appliance is the dryer. Disaggregation of this appliance was attempted in the aggregate signal of house-1 from the REDD dataset in order to test a certain aspect of the designed technique, the ability to disaggregate a model of the appliance even when some of the main changes of the profile do not occur at all during its operation.

In figure 9.3 is presented the signature of the dryer in house-1 of the REDD dataset, in comparison to a signature from house-3 of the REDD dataset. The first is the only model of the dryer in the datasets that the low-power states are not observed. The profile created by the Profiler module, after processing all the dryer models, represents the last of these low power states with one of the main blocks (see figure 6.10c). In order to allow the detection of appliances that do not present all the states during their operation, when the Pattern Recognition module creates the various candidates out of the possible combinations of the detected power changes, additional candidates that are missing one or more main changes are created (see section 7.1). This functionality is tested with the disaggregation of the particular model of dryer.

Unlike the cases of the dishwasher and the washing machine, for this evaluation the classifier had to be trained on the particular model of dryer that would be disaggregated. The reason is that it is the only model with this kind of signature and the classifier needs to be trained that there are dryers that do not present all the states during their operation. The classifier was trained on eight signatures of the model. The disaggregation was performed on aggregate data of 20 days length, during which 10 operations of the appliance occured, different than the ones trained on. The results are presented in table 6.

Table 6: Dryer disaggregation: Results

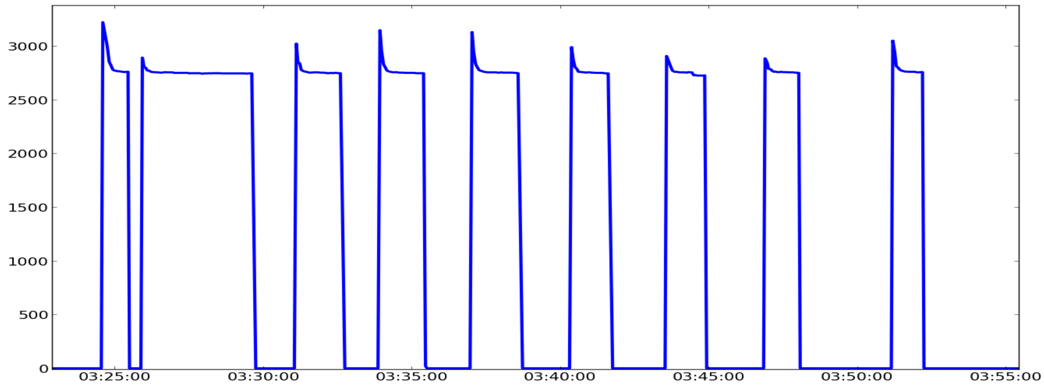
	REDD:House-1
Actual Number of Signatures	10
Actual Power Consumption	11.56 kWh
Number of Signatures Disaggregated	6
Number of True Positives	5
Number of False Positives	1
Energy Consumption Correctly Disaggregated	5.52 kWh
Disaggregation Accuracy	50%
Detection Accuracy	83.3%
Percent of Energy Correctly Disaggregated	47.8%
Percent of Energy Disaggregated from Detected Events	85.8%

In this test case the disaggregation accuracy is fairly lower than for the previous cases. However, the detection accuracy remains high, with only one false positive detected. The percent of energy that the system manages to disaggregated is also satisfying, like in the previous cases.

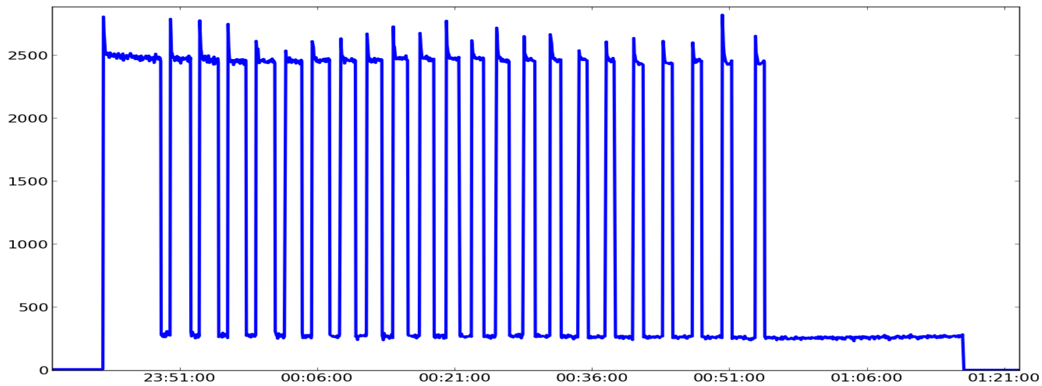
Inefficiencies

The disaggregation of the dryer proved quite more inaccurate in comparison to the previous cases, even though the classifier was trained on the particular dryer model. There are two main reasons for this inefficiency. The first one is the fact that this particular model of a dryer is different than the rest. The classifier is trained on all of them, but is less confident in disaggregating this particular model since most of its training samples are different.

The second reason is the problematic detection of the multiple heating cycles, which are the dryer's most distinctive feature. Those cycles are modelled in the appliance's profile as repeats



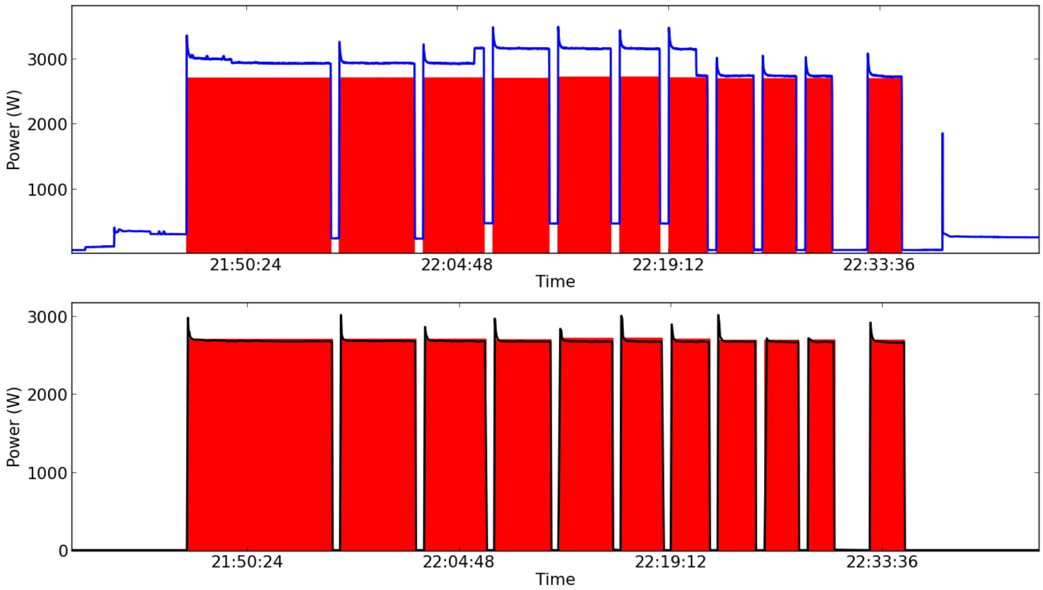
(a)



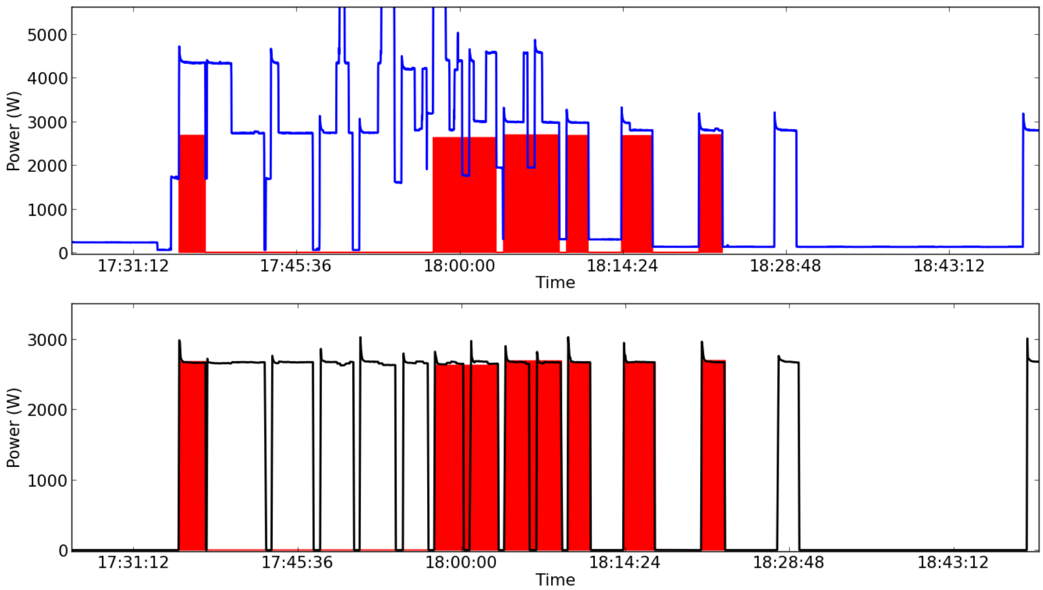
(b)

Figure 9.3: Signatures of two different models of dryers. During the operation of the dryer from house-1 of the REDD dataset (a) the low-power states are not observed. In (b) is presented a dryer's signature from house-3 of the REDD dataset. The shape of this signature is the most commonly observed among the signatures of the various dryer models in the available datasets.

of the first main block. Their detection, one of the latest extensions to the system, is performed in a rather simplistic way. Their detection is mainly based on finding power changes with similar magnitudes in a certain area specified by the profile, with fairly consistent distances from each other. In occasions where multiple such repeating cycles occur with durations and distances that vary, the detection mechanism is inefficient. Inaccurate detection of such parts of a signature are depicted in figure 9.4b), the false positive detected event.



(a)



(b)

Figure 9.4: Two disaggregated events during the evaluation of the dryer. An accurately disaggregated operation of the dryer is depicted in (a). The detection of the repeating parts of the dryer's operation is only based on detecting power changes with similar magnitudes and at fairly stable distances from each other. The detection mechanism proves to be inefficient in occasions that the distances between some of the parts vary significantly, such as the operation of the dryer depicted in (b). In addition, intermediate power changes are observed, which bring the power below the height of the detected blocks, which is an obvious inefficiency of the current version of the detection mechanism for the repeating parts.

10 Conclusions and Future Work

10.1 Achievements

The designed system has shown very promising results, even though the disaggregation accuracy was not exceptional in all the cases of the evaluation. First of all, the system was able to **form a generic profile** for each of the targeted multi-state appliances after processing the signatures from the various models, by detecting their main states and power changes. The rough shape of an appliance's signature is 'mapped' by the use of the distances between the main power changes, as well as their magnitudes. Various features extracted from the signatures of the targeted appliances complement the profile. With the use of Gaussian Mixture Models, it has been achieved to incorporate density functions to the profile that describe the values that the various features take for the different models of an appliance. During the evaluation, the system was **able to generalize and detect unknown models** of a dishwasher and washing machine that were not processed during the creation of the profile with satisfying accuracy. The evaluation for the dryer showed that the system is also able to detect models of an appliance that might not present certain main power changes of the profile during their operation, albeit with lower accuracy.

The results of the evaluation display **high detection accuracy** of the system. In every tested case, the number of false positives detected in the aggregate signal was small. This is a result of the use of a large number of features for forming an appliance's profile. The profile is able to describe the various characteristics of an appliance's operation, allowing the disaggregation of true operations. On the other hand, the various random patterns in the aggregate signal might present some similarities with an actual operation of the disaggregated appliance, but it is unlikely to closely match the profile in every aspect. As a result, the classifier detects the missing features as anomalies and the case of false positives is generally avoided. This is one of the main attributes of the designed technique. The high detection accuracy of the technique can be a stepping stone for further work, building on top of the high confidence that the detected events are true operations and not false positives.

The system is fairly **successful in disaggregating the energy consumption of a detected event**. The profile formed for an appliance describes the rough shape of an appliance's signature by determining the main parts that the signature consists of. Parts of the appliance's operation that are not observed consistently during the operation of the various models are not added to the profile. Subsequently, they are not detected. However, for the cases of the three targeted appliances, the undetected parts of the operation usually consume low amounts of power or their durations are fairly short. Subsequently the energy consumed during those parts of the operation is low and the overall percentage of disaggregated energy is not severely reduced.

10.2 Limitations

The investigated approach and the implemented system do have their limitations. The technique for forming the generic profile of an appliance is based on detecting the states that are consistently observed during the operation of the various models. This sets the requirement that **the signatures of the various appliance models need to present similarities** in the first place. For instance it is impossible for a single profile to be formed that describes both the top-loader washers (see section 3.2.1) and the washing machine type focused in this work. Other such problematic cases can be appliances that their signatures vary significantly, depending on the selected program of operation. Such a case is the microwave. Depending on the program selected, its operation may last half a minute or well over ten minutes. The profile formed would not be able to model such behaviour accurately.

Another problem regards the generalization capability of the technique, the ability to detect models different than the ones processed during the profiling process and the training. The greatest limitation in this regard has been introduced by the design of the pattern recognition technique. As explained in section 7.6, the system **fails to detect unknown appliance models that during the main states of their operation consume lower or higher power than the minimum and maximum values observed in the signatures of the various models during the profiling process**. The same limitation applies for the distances between the main power changes during their operation. For instance, if the system was called to disaggregate a model of a washing machine that is fairly bigger and more power hungry than the models in the available datasets, the attempt would fail. However this limitation is not of this approach alone. Most supervised pattern recognition techniques have to assume a fairly complete training set, that is representative of the whole appliance type.

Another problematic point is that **parts of an appliance's operation that are not observed in the majority of the models' signatures are not modelled by the profile**. As a result they are undetectable during the pattern recognition phase. As explained above, in the case of the appliances targeted during this work, those non-profiled parts do not contribute much to the total energy consumption of the appliance. In the cases of other appliances, however, perhaps this loss would be greater. Moreover, it is very important for a disaggregation technique to be able to 'clear' the detected event from the aggregate signal. By doing so, it is possible to use the technique in combination with another algorithm, that could detect the rest of the appliances. The current version of the system, however, does not detect all the power changes generated by the disaggregated appliances. Removing the detected changes from the aggregate signal would still leave a number of changes behind. These changes could be mistakenly detected by another disaggregation system as operations of another appliance.

The current version of the system also has the limitation that it **cannot disaggregate two appliances of the same type if they operate simultaneously**. As explained in section 7.4, in a part of the aggregate signal that an actual operation of a multi-state appliance occurs, there can be detected various candidate-patterns resembling the 'rough shape' of the appliance's signature, as described by the profile. Out of those candidates only one is disaggregated, the one with the highest score, as long as it is classified positively by the classifier. As a result, if for instance two washing machines in a laundry room operate at the same time, only one will be disaggregated. An extension to the system could give the solution to this problem, by performing a second scan of the aggregate signal after the first detected instances were removed from the signal.

Finally, the technique **would fail to create a profile and detect appliances which do not present discrete power levels** during their operation. Examples of such continuously varying loads are dimmer lights and motors controlled by adjustable speed drives.

10.3 Future Work

Learn the particular model in the aggregate and perform a second-pass disaggregation

The designed technique has presented a high detection accuracy, avoiding the detection of false positives. Even if some operations of an appliance are missed during the disaggregation, it can be inferred that most of the detected events are true positives. Based on this, it could be possible to extract the characteristics of the particular model of the appliance that is present in the house by processing the disaggregated events. For instance, it is possible to extract the actual amount of power it consumes in the various states, as well as the durations of the states. In addition, by processing the aggregate signal in the areas that the events were detected, patterns that are created by the particular model's operation could be extracted since they would consistently appear in those areas of the aggregate. Using the new information, a new and more accurate profile could be formed, adapted to the characteristics of the particular model. Finally, a second-pass over the aggregate signal could be performed. The second pass could find operations of the appliance that were missed during the first pass, re-evaluate the confidence of the already detected events, or even detect parts of the appliance's signatures that were missed in the first pass because they were not part of the generic profile.

Combine the designed technique with different algorithms

The designed system could be used in combination with other techniques, in order to complement each other. For instance, it could be used in combination with an algorithm that is more accurate in disaggregating the whole signature of an appliance, even if it is more prone to false positives. Since our designed technique presents high detection accuracy, it could report the areas of disaggregated events to the second algorithm. The latter would only scan the specified areas. This way the false positives are still avoided and the weakness of the designed system, leaving out parts of a signature, is countered.

Further investigation for features and their addition to the profile

The use of a big number of features for the detection of an appliance seems to be promising. Further investigation could be performed to extract more patterns that may help distinguish an appliance. Such features could be the magnitude of a spike's overshoot or the power decays that follow them in some cases. To take it even further, the investigated features could be of non-power nature. For instance, it could be investigated whether certain appliances are usually operated at certain parts of the day. Perhaps dishwashers tend to be operated the next hours after lunch or dinner time. The designed technique has a very convenient attribute; it is fairly easy to extend in order to incorporate any additional feature into the profile of an appliance. Just an additional entry to the profile and an additional score assigned to the candidates will suffice.

Addition of less consistent features into the profile.

In the current version of the system, a feature is added to the generic profile of an appliance only if it is observed in the majority of the signatures of the various models. However, it would be possible to extend the technique in order to add less consistent features to the profile with an appropriate use of weights. Such an extension would improve the accuracy when disaggregating models that exhibit such features. In addition, the percentage of energy disaggregated would be improved, by enabling the detection of parts of the signatures that are currently left out.

Optimization of the pattern recognition process

Throughout the development of the presented system, one of the main problems was to avoid making the process too computational heavy. As explained in section 7.5, a big number of candidate instances are generated during the pattern recognition process. Many attempts and extensions were made to reduce that number, gradually making the process far more optimized than the first versions of the system. Still, there is a lot of room for improvement. For instance, after every evaluation of a feature, candidates with very low scores assigned for that feature could be carefully removed, making the rest of the evaluation lighter. Such an improvement would not only affect the disaggregation time. As explained in the end of subsection 6.2, the main reason that only three main blocks are chosen for creating the appliance's profile is exactly this computational inefficiency. An improvement over it would allow for the choice of one more main block and subsequently describe in more detail the operation of the appliance.

Investigate more appliances

The three appliances that were targeted in this work were chosen after an initial investigation. Their signatures are complicated but the various models consistently present certain patterns which enable their profiling as a 'set of features'. More appliances should be investigated in order to find such consistent patterns, extend the system and take the approach further.

Appendices

A Form of the Generated Profile

The creation of the profile is described in section 6. The form of the various entries of the profile are presented below:

The entry of the profile for the magnitudes of the main power changes consists of a sub entry for each main change, consisting of the min value observed, the max value observed, the parameters of the Gaussian Mixture representing the density function of the magnitudes observed in the signatures and, finally, the weight of the main change, which is the percentage of the signatures that the main change actually existed. The form of the entry is as follows:

```
[
[min_magn_0, max_magn_0, [parameters_of_the_gmm_0], weight0],
. . . . . . . . . . . . .  . . . . . . . . . . . . . . . . .
 [[min_magn_5, max_magn_5, [parameters_of_the_gmm_5], weight5]
]
```

If correlations between the magnitudes of the certain main changes were observed, an entry is created to the profile to describe it. For instance, if a correlation has been found for changes X,Y and Z, the following entry describes that the magnitude of X must be equal to the magnitude of Y, or equal to the addition of Y and Z:

```
[main_change_X, main_change_Y, main_change_Z]
```

The entry for the distances between the main changes is similar to the entry for the magnitudes. However it consists of one sub entry for each main change which describes the distances of all the other main changes from this particular one. For instance, the sub entry for the main-change-0 is as follows:

```
[0,
[min_dist0_1, max_dist0_1, [parameters_of_the_gmm_0_0]],
. . . . . . . . . . . . . . . . . . . . . . . . . . . . .
[min_dist0_5, max_dist0_5, [parameters_of_the_gmm_0_5]]
]
```

In the case that repeating parts of the signature were found, an entry is added to the profile with the type of the repeating parts, the main change with which the repeating parts have a similar magnitude, and two further sublists. The type of the repeating parts entry can be one of the following:

- '1only' - meaning that only one repeat of the cycle was observed consistently in the
- signatures.
- '1or2' - meaning that in most of the signatures there were either one or two repeats of the cycle.
- 'consistent' - meaning there were at least three repeats in most of the signatures and their duration and the time intervals between them were fairly stable.
- 'any-number' - describes the case that there have been repeated parts observed in most of the signatures but their number is fairly random.

These two sublists describe the area in which the repeating parts have been observed. By providing the number of a main change of reference and the minimum and maximum distance from it, the beginning and the end of the area that the repeating parts have been observed is defined. The entry is of the form:

```
[type, main_change_with_similar_magnitude,
[main_change_X, min_distance_from_X, max_distance_from_X],
[main_change_Y, min_distance_from_Y, max_distance_from_Y],
]
```

The type of each main change is described by a list of percentages, one for each main change. Those percentages represent the percentage of the signatures that the corresponding main change was found to be a spike. The entry is a simple list of the form:

```
[percentage_0, . . . , percentage_5 ]
```

The existence of rapid changes is described by an entry with two sublists. The first one contains the minimum and maximum distances that the periods of rapid changes were observed to begin from a certain main change. The second sublist contains the same information for the point that the rapids are ending. The main changes of reference are also included, which are chosen to be the main changes closest to the beginning and ending points of the periods of rapids. The entry is of the form:

```
[[main_ch_X, min_dist_from_X_raps_start, max_dist_from_X_raps_start],
 [main_ch_Y, min_dist_from_Y_raps_end, max_dist_from_Y_raps_end]]
```

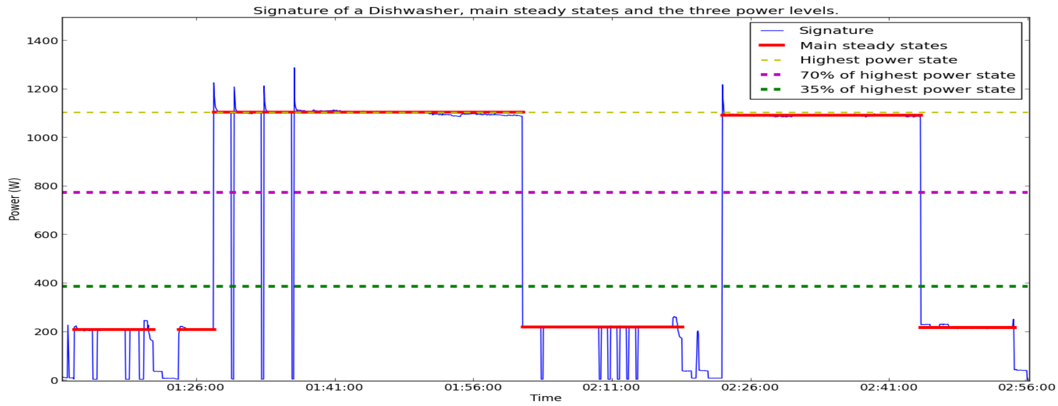
An entry of the following form describes that the ratio between the duration of X and the duration of Y has been found to have a density function modelled by the given GMM function:

```
[block_X, block_Y, [parameters_of_the_gmm]]
```

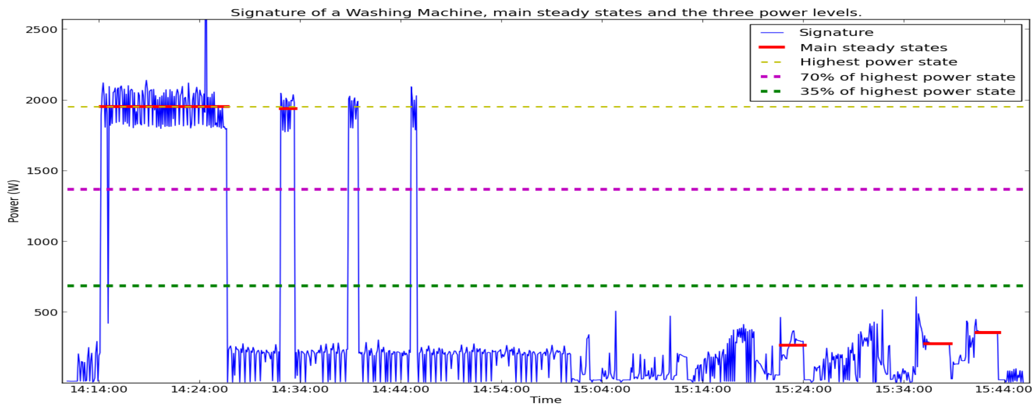
B Diagrams

B.1 Forming An Appliance's Profile

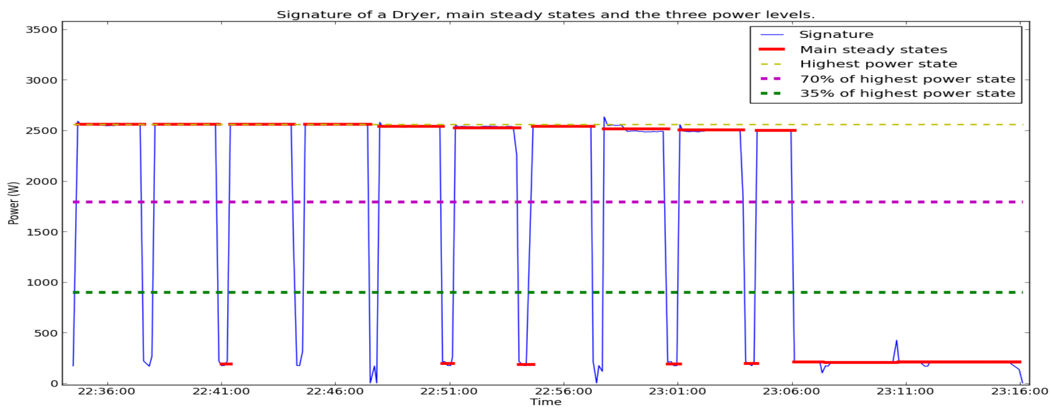
B.1.1 Separation of Steady States into Power-Groups



(a)



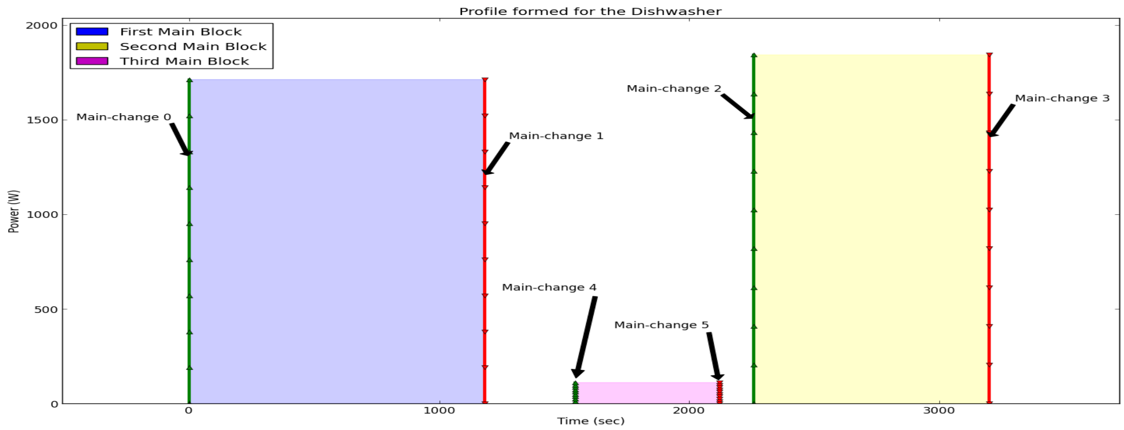
(b)



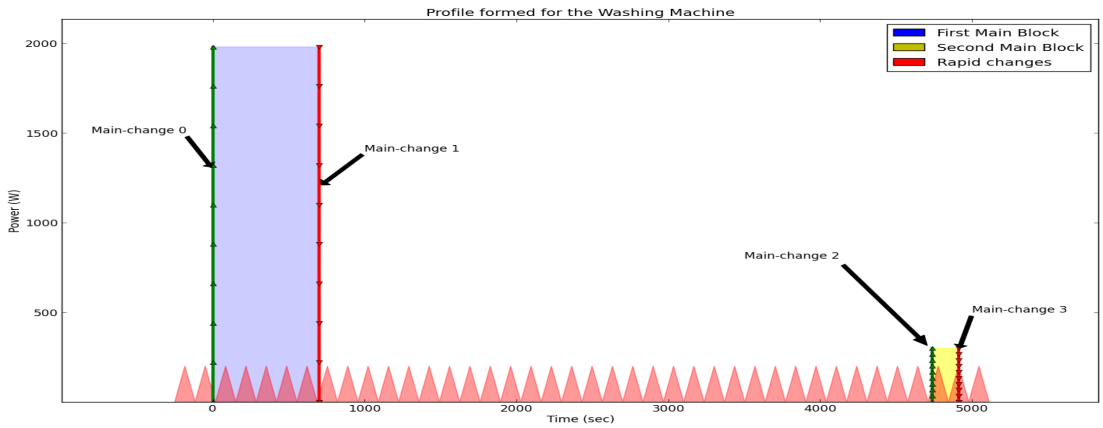
(c)

Figure B.1: Signatures of (a) dishwasher, (b) washing machine, (c) dryer, the steady states extracted by the Mapper and their separation into three power-groups in the first step of choosing the main blocks of an appliance's generic profile (section 6.2).

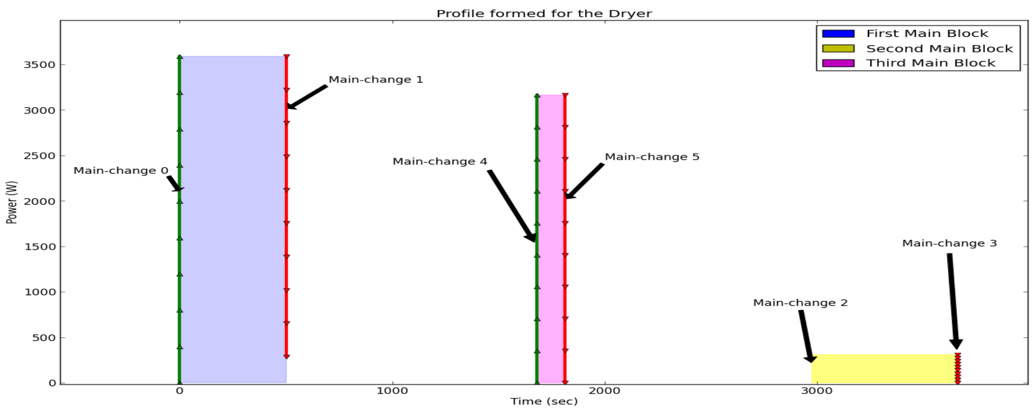
B.1.2 Chosen Main Blocks and Main Changes of the profiles



(a)



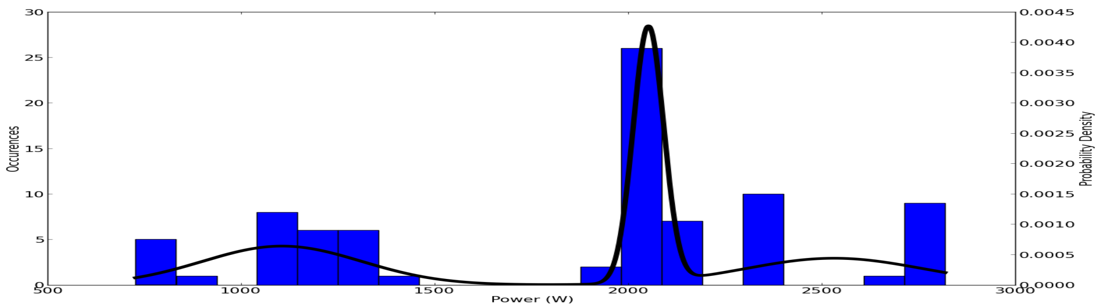
(b)



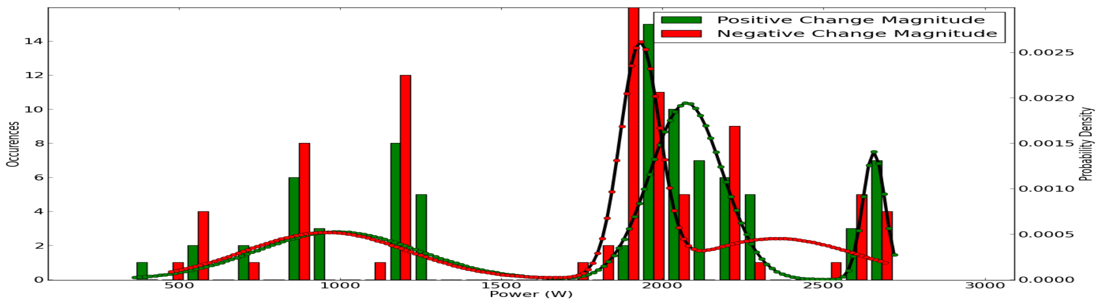
(c)

Figure B.2: The main blocks and the main changes of the profile generated for (a) the dishwasher, (b) washing machine, (c) dryer. Up to three main blocks are chosen by the algorithm described in section 6.2, in order to represent the rough shape of the signatures. The magnitudes and distances between the main changes depicted are the means of the values found while processing the various signatures and are used only to help visualize the blocks' positioning. The actual profile incorporates Gaussian mixture functions to describe the values observed in the signatures.

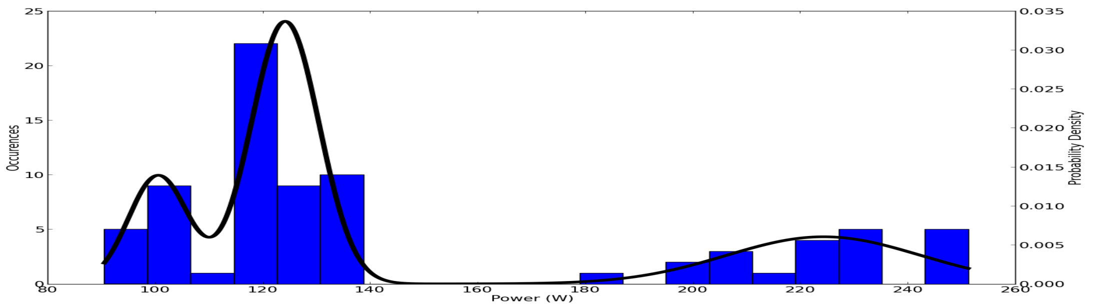
B.1.3 Power Consumption and Magnitudes of Main Changes



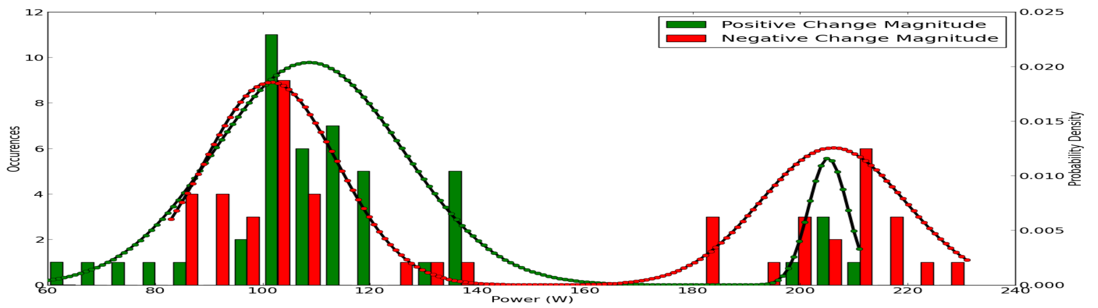
(a)



(b)

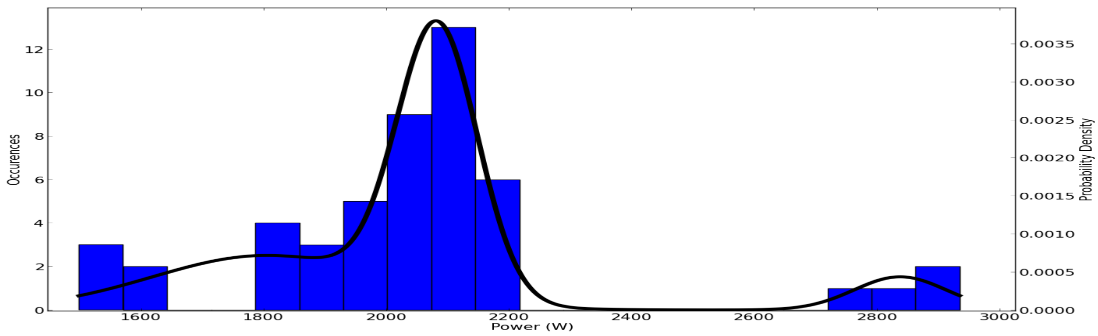


(c)

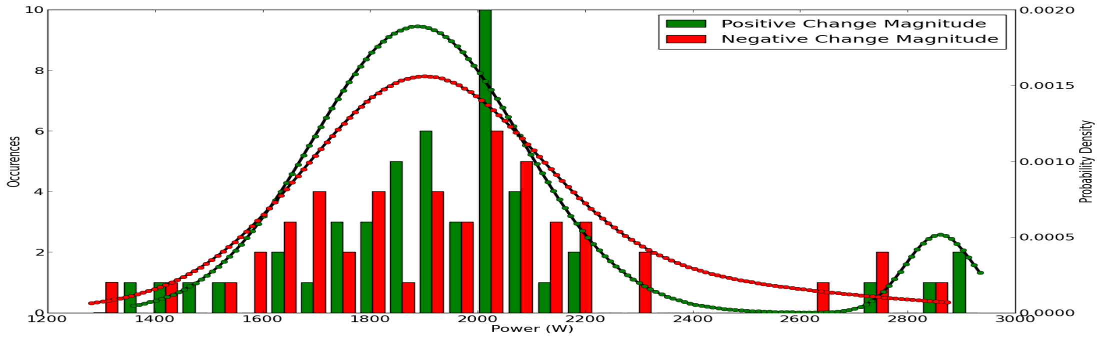


(d)

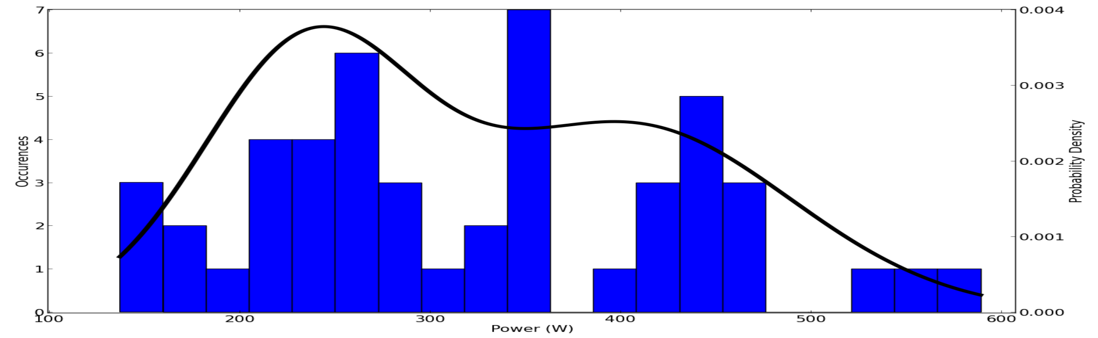
Figure B.3: (a) Histogram and density function of the power consumption observed during the high power states of dishwashers. (b) Magnitudes of the positive and negative power changes of the first chosen block (first heating cycle) of the dishwasher profile. (c) Histogram and density function of the power consumption observed during the low power states of dishwashers. (d) Magnitudes of the positive and negative power changes of the third (low power) chosen block of the dishwasher profile.



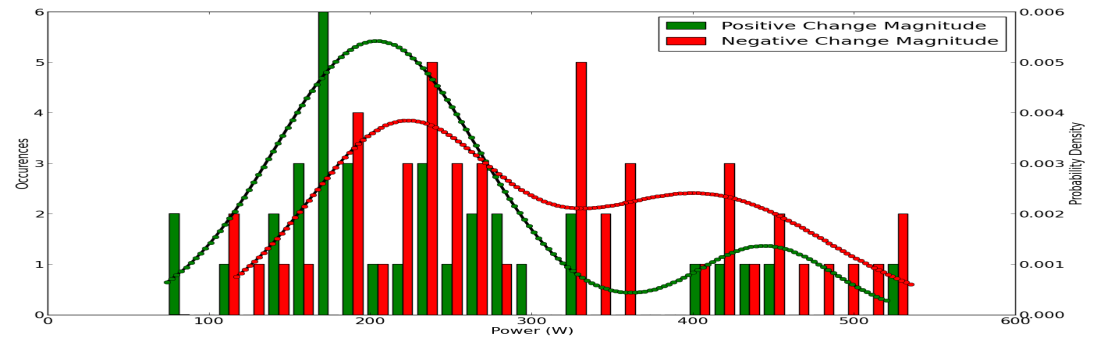
(a)



(b)

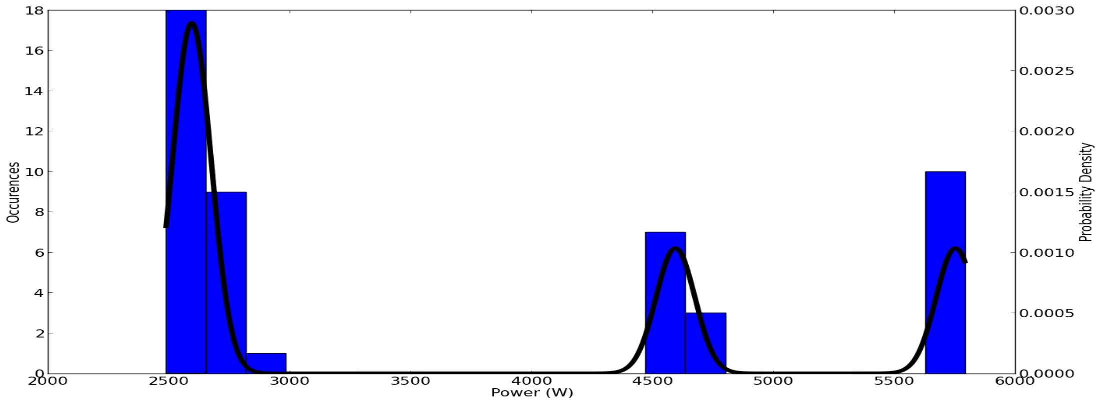


(c)

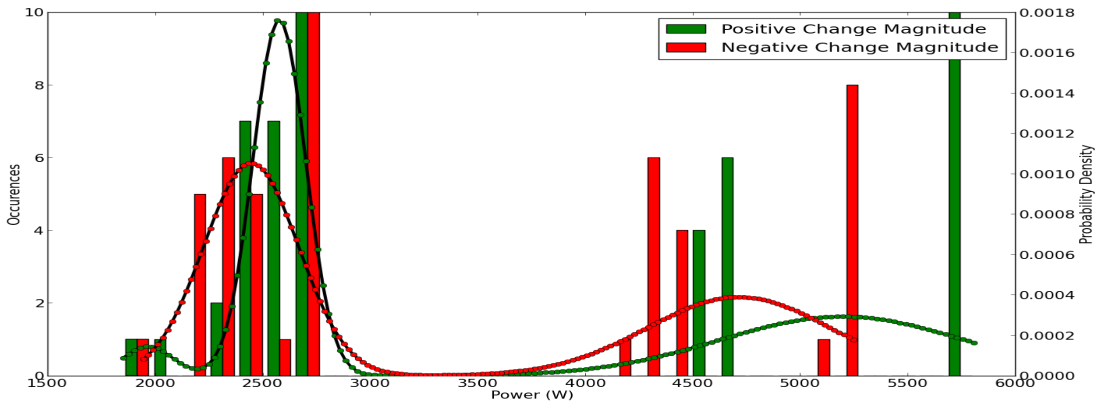


(d)

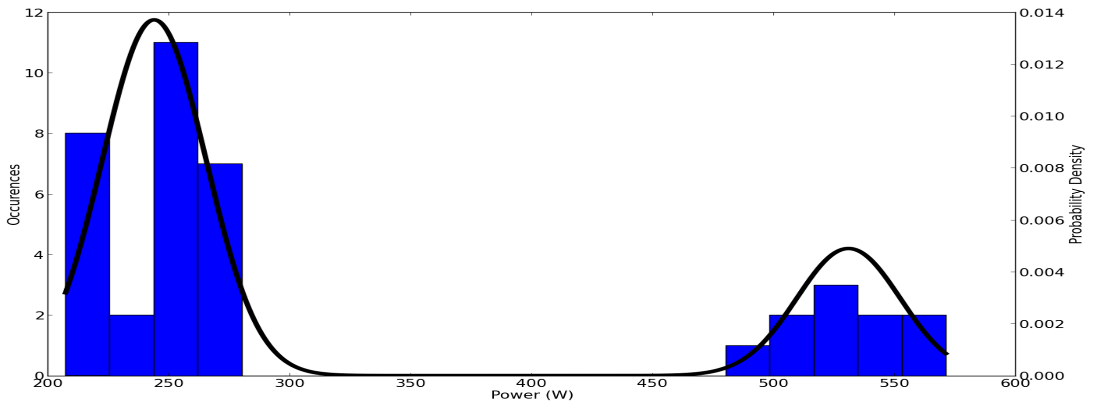
Figure B.4: (a) Histogram and density function of the power consumption observed during the high power states of washing machines. (b) Magnitudes of the positive and negative power changes of the first chosen block (heating cycle) of the washing machine profile. (c) Histogram and density function of the power consumption observed during the low power states of washing machines. (d) Magnitudes of the positive and negative power changes of the second (low power) chosen block of the washing machine profile.



(a)



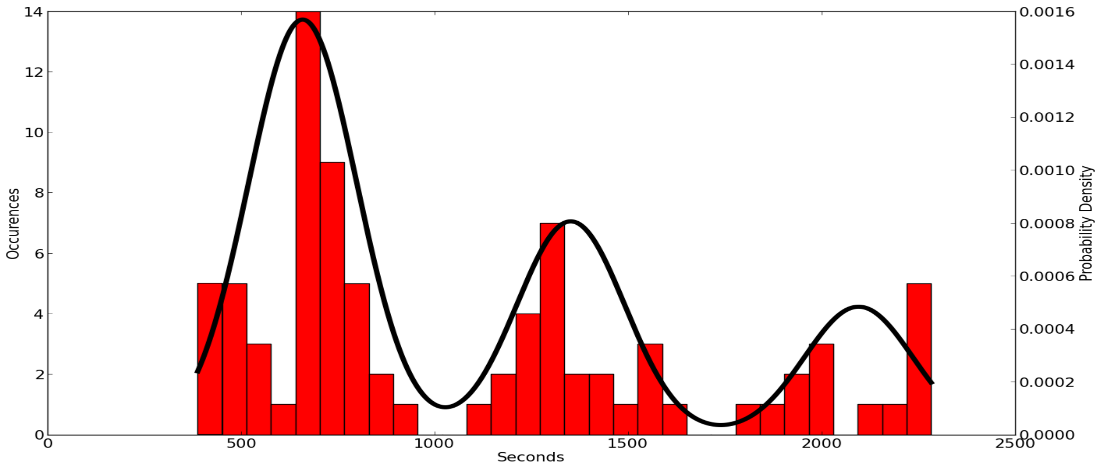
(b)



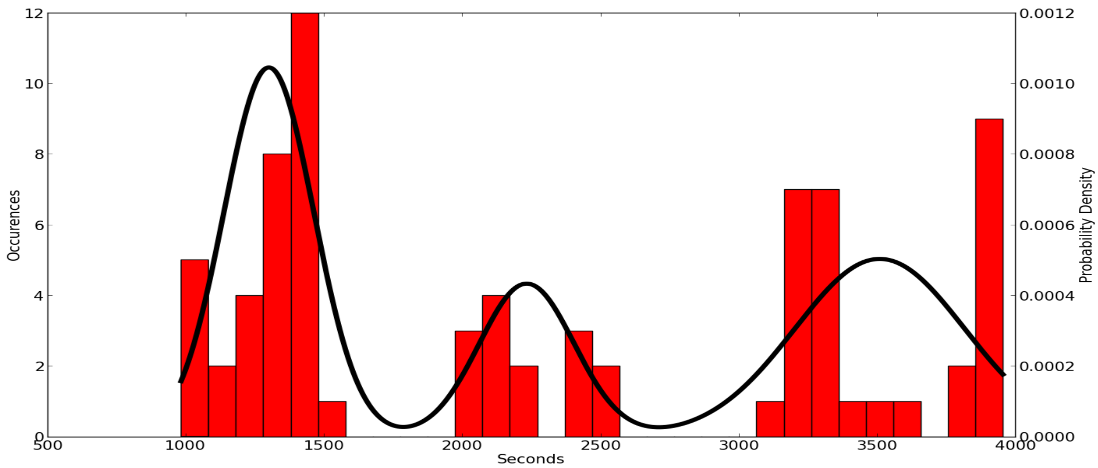
(c)

Figure B.5: (a) Histogram and density function of the power consumption observed during the high power states of dryer. (b) Magnitudes of the positive and negative power changes of the first chosen block (heating cycle) of the dryers. (c) Histogram and density function of the power consumption observed during the low power states of dryers.

B.1.4 Durations of States and Distances of Main Changes

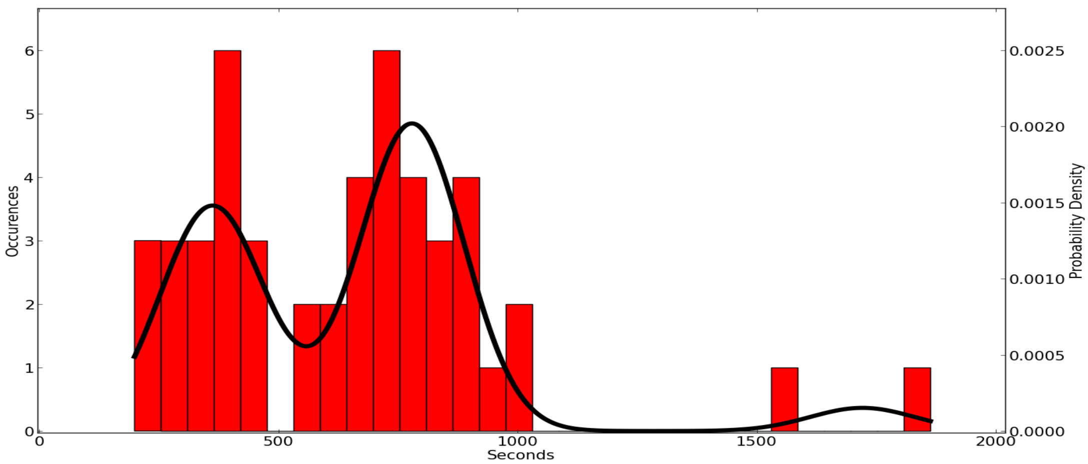


(a)

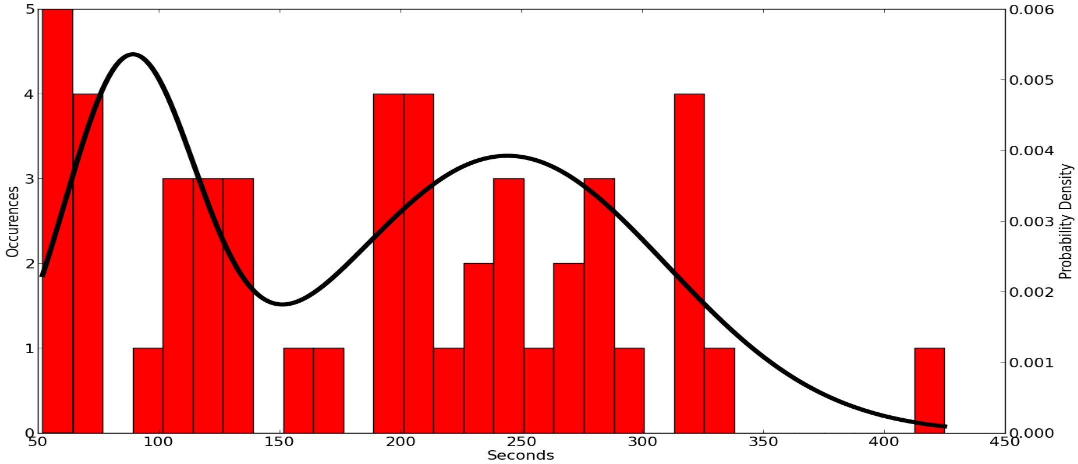


(b)

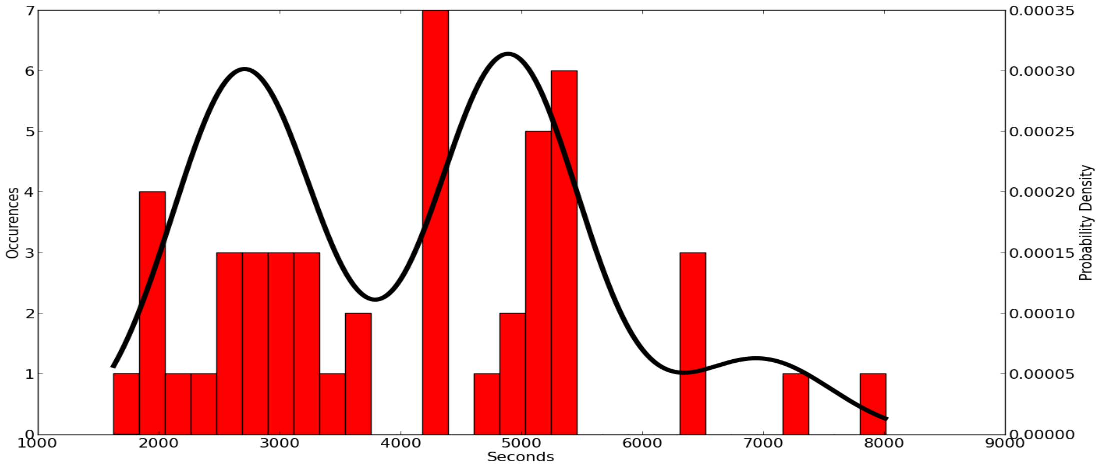
Figure B.6: (a) Histogram and density function of the duration of the first heating cycle, as observed in the various signatures of dishwashers. (b) Histogram and density function of the time interval between the beginning of the two heating cycles of dishwashers.



(a)

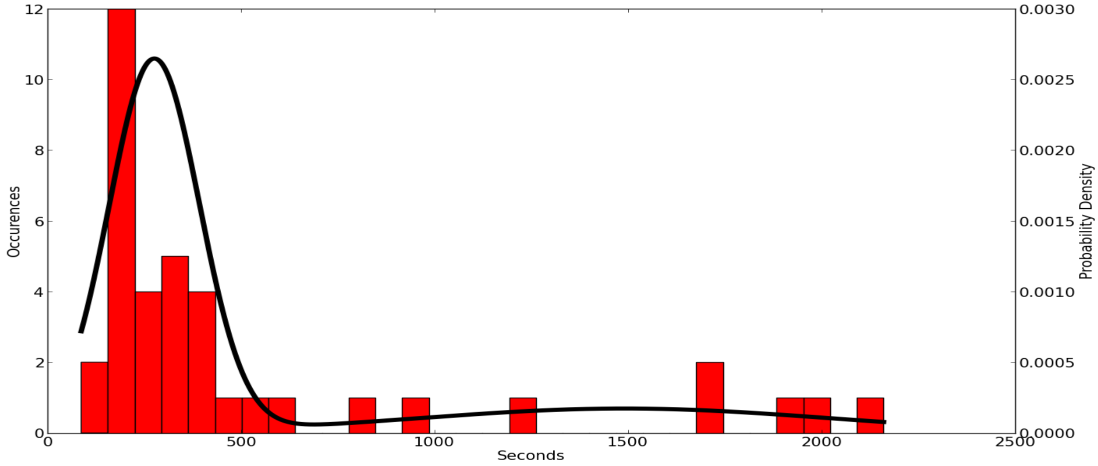


(b)

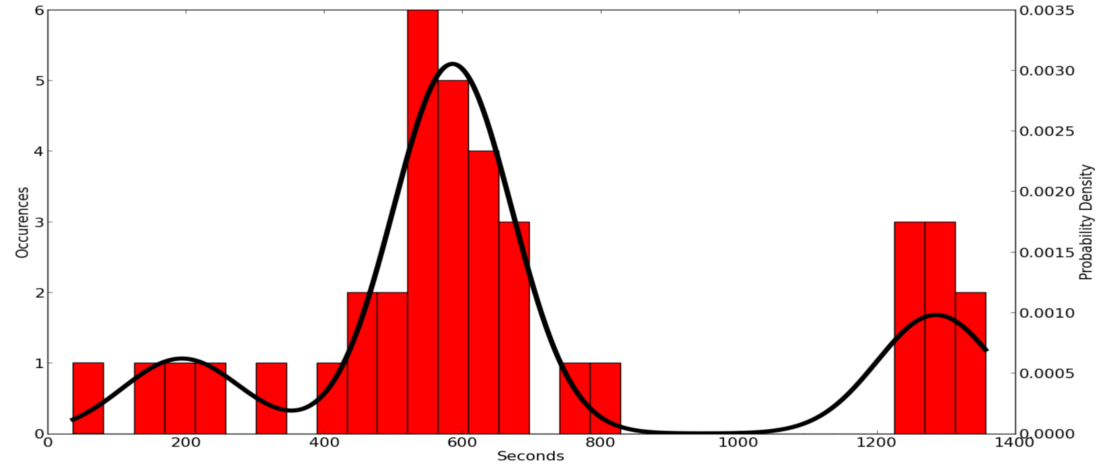


(c)

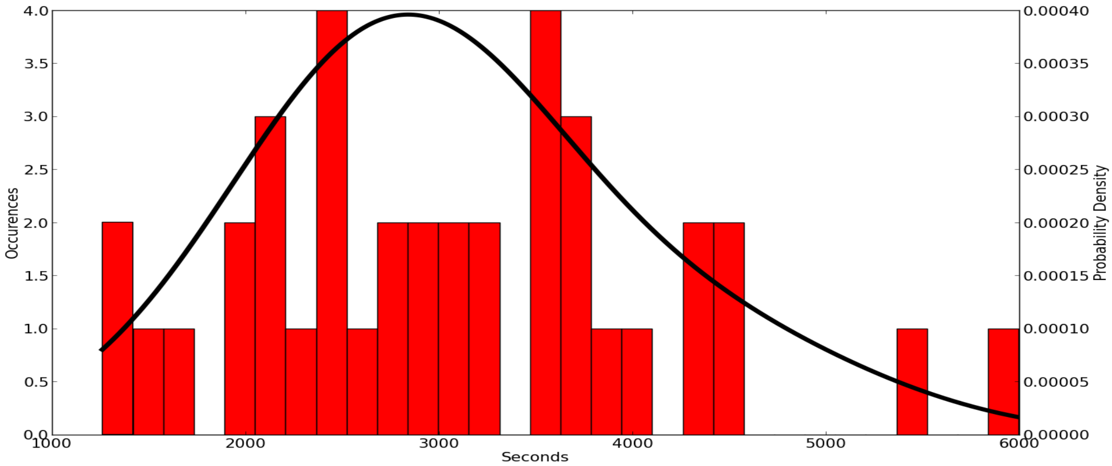
Figure B.7: Histograms and density functions of (a) the duration of the first chosen block (heating cycle), (b) second chosen block, (c) the time interval between the beginning of the two blocks, as observed in signatures of washing machines



(a)



(b)



(c)

Figure B.8: Histograms and density functions of (a) the duration of the first chosen block, (b) second chosen block, (c) the time interval between the beginning of the two blocks, as observed in signatures of dryers

B.1.5 Repeating Parts of the Signatures

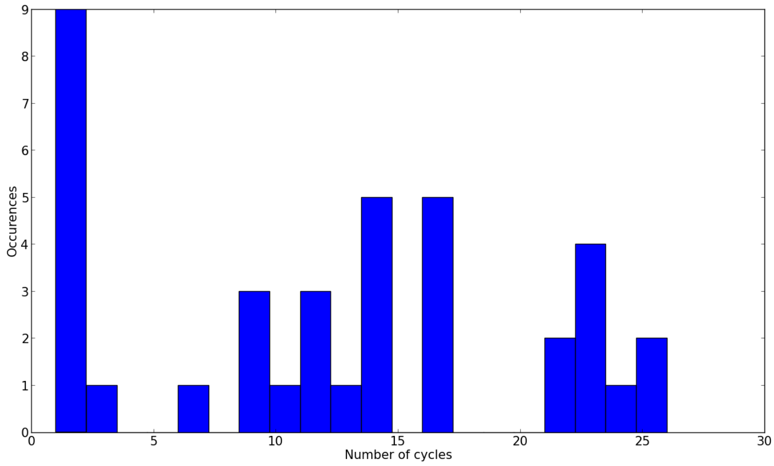


Figure B.9: Histogram of the number of heating cycles found in the signatures of dryers. The heating element turns on and off several times for most models of dryers.

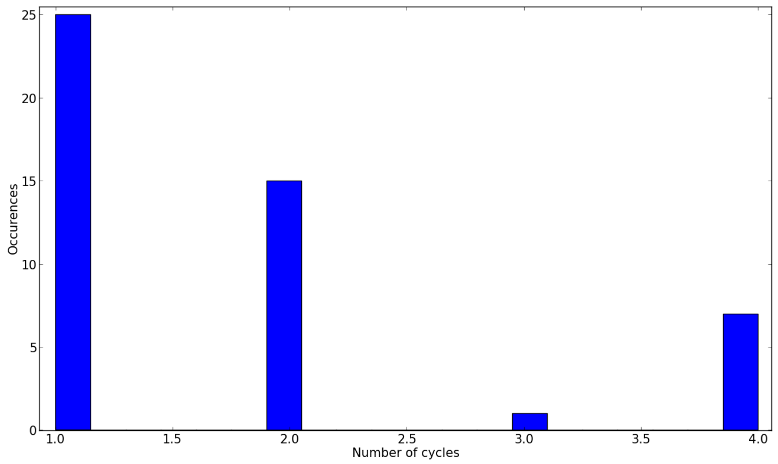


Figure B.10: Histogram of the number of heating cycles found in the signatures of washing machines. Just like the dryer, during the operation of many models the heating element turns on and off more than once.

11 Bibliography

References

[1] Department of Energy and Climate Change , *"UK Energy in Brief 2012"*. National Statistics Publication, July 2012.

[2] Roy Dong, Lillian Ratliff, Henrik Ohlsson and S. Shankar Sastry, *A Dynamical System Approach to Energy Disaggregation*. 52nd IEEE Conference on Decision and Control, April 2013.

[3] Department of Energy and Climate Change , *"Energy Consumption in the UK"*. 26 July 2012.

[4] Paul Bolton , *"Energy Prices Commons Library Standard Note"*. March 2013.

[5] Department of Energy and Climatic Change and Ofgem , *"Smart Metering Implementation Program Prospectus"*. Ofgem and Ofgem E-Serve, July 2010.

[6] Department of Energy and Climatic Change and Ofgem , Smart Metering Implementation Program Update, DECC, April 2012.

[7] Pike Research (renamed to Navigant Research), *"Smart Meters"*. Quarter2 of 2012.

[8] Energy Information Administration , *"Annual Energy Outlook 2009, With Projections to 2030"*. p.71, March 2009.

[9] Sarah Darby , *"The effectiveness of feedback on energy consumption, A review for Defra of the literature on metering, billing and direct displays"*. Change Institute, University of Oxford, April 2006.

[10] Corinna Fischer , *"Feedback on household electricity consumption"*. Springer Science and Business Media B.V., May 2008.

[11] Karen Ehrhardt-Martinex, Kat A. Donnely & John A. Skip Laitner , *"Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household electricity-Saving Opportunities"*. ACEEE, June 2010.

[12] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert , *"Is disaggregation the holy grain of energy efficiency?"*. Energy Policy, vol. 52, January 2013.

[13] Jack Kelly , *"Disaggregating Smart Meters Readings using Device Signatures"*. Imperial College London -Department of Computing, September 2011.

[14] M. Zeifman and K. Roth , *"Nonintrusive Appliance Load Monitoring: Review and Outlook"*. IEEE transactions on Consumer Electronics, February 2011.

[15] M. Zeifman and K. Roth , *"Nonintrusive Appliance Load Monitoring: Review and Outlook"*. International Conference on Consumer Electronics, January 2011.

[16] A. Zoha, A. Gluhak, M. Ali Imran and S. Rajasegarar , *"Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey"*. Center of Communication Systems Research - University of Surrey, Department of EEE - University of Melbourne, December 2012.

[17] G.W. Hart, *"Nonintrusive Appliance Load Monitoring"*. Proceedings of the IEEE, December 1992.

[18] A.G. Ruzzelli, C. Nicolas, A. Schoofs and G.M.P OHare, *"Real-Time Recognition and Profiling of Appliances through a Single Electricity Sensor"*. Sensor Mesh and Ad Hoc Communications and networks (SECON), 2010 7th Annual IEEE Communications Society Conference, June 2010.

[19] L. Farinaccio and R. Zmeureanu, *"Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses"*. Energy and Buildings, February 1999.

[20] M. Baranski and J. Voss, *"Genetic Algorithm for Pattern Detection in NIALM Systems"*. IEEE International Conference on Systems, Man and Cybernetics, 2004.

[21] J. Liang, S.K.K. Ng, G. Kendall, J.W.M. Cheng, *"Load Signature Study-Part I: Basic Concept, Structure and Methodology"*. IEEE Transactions on Power Delivery, April 2010.

[22] J. Liang, S.K.K. Ng, G. Kendall, J.W.M. Cheng, *"Load Signature Study-Part II: Disaggregation Framework, Simulation, and Applications"*. IEEE Transactions on Power Delivery, Aprill 2010.

[23] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford and P. Armstrong, *"Power Signature Analysis"*. IEEE power & energy magazine, March/April 2003.

[24] Leslie K. Norford and Steven B. Leeb, *"Non-intrusive electrical load monitoring in commerical buildings based on steady-state and transient load-detection algorithms"*. MIT, September 1995.

[25] S. N. Patel, Thomas Robertson, Julie A. Kientz, Matthew S. Reynolds, and Gregory D. Abowd, *"At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line"*. Conference on Ubiquitous Computing, 2007.

[26] H.H. Chang, C.L. Lin and J.K. Lee, *"Load Identification in Nonintrusive Load Monitoring Using Steady-State and Turn-on Transient Energy"*. Proceeding of the 2010 14th International Conference on Computer Supported Cooperative Work in Design, April 2010.

[27] J. Z. Kolter and M.J.Johnson, *"REDD: A Public Data Set for Energy Disaggregation Research"*. SustKDD 2011 San Diego, August 2011.

[28] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, Ralf Steinmetz, *"On the Accuracy of Appliance Identification Based on Distributed Load Metering Data"*. Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT), pp. 1-9, 2012.

[29] Sean Barker, Aditya Mishra, David Irwin, Emmanuel Cecchet, Prashant Shenoy, and Jeannie Albrecht, *"Smart*: An Open Data Set and Tools for Enabling Research in Sustainable Homes"*. Proceedings of the 2012 Workshop on Data Mining Applications in Sustainability (SustKDD 2012), Beijing, China, August 2012.

[30] Stephen Makonin, Fred Popowich, Lyn Bartram, Bob Gill, and Ivan V. Bajic, *"AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research"* in Electrical Power and Energy Conference (EPEC), 2013 IEEE, pp. 1-6, 2013.

[31] R. Dong, L Ratliff, H. Ohlsson and S.S. Sastry, *"A Dynamical Systems Approach to Energy Disaggregation"*. 52th IEEE Conference on Decision and Control, April 2013.

[32] K.C. Armel, A. Gupta, G. Shrimali, A. Albert, *"Is disaggregation the holy grail of energy efficiency? The case of Electricity"*. Energy Policy, January 2013.

[33] H. Kim, M. Marwah, M. Arlitt, G. Lyon and J. Han, *"Unsupervised Disaggregation of Low Frequency Power Measurements"*. 11th International Conference on Data Mining, 2011.

[34] A. I. Cole and A. Albicki, *"Algorithm for nonintrusive identifucation of residential appliances"*, in Proceedings of the 1998 IEEE International Symposium on Circuits and Systems, 1998, June 1998.

[35] A. I. Cole and A. Albicki, *"Data extraction for effective non-intrusive identification of residentail power loads"*, in IEEE Instrumentation and Measurement Technology Conference, 1998, May 1998.

[36] Pulak Purkait, *"Optical Handwritten Character/Numeral Recognition"*, Indian Statistical Institute, 9th North-East Workshop on Computational Information Processing.

[37] Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, Ching Y. Suen, *"Character Recognition Systems, A Guide for Students and Practitioners"*, Wiley-Interscience publication.

[38] Juan-Carlos Perez, Enrique Vidal, Lourdes Sanchez, *"Simple and Effective Feature Extraction for Optical Character Recognition"*, 5th Spanish Symposium on Pattern recognition and images analysis, 1994.

[39] Douglas Reynolds, *"Gaussian Mixture Models"*, MIT Lincloln Laboratory.

[40] In Jae Myung, *"Tutorial on maximum likelihood estimation"*, Journal of Mathematical Psychology, 2001

[41] Frank Dellaert, *"The Expectaion Maximization Algorithm"*, College of Computing, Georgia Institute of Technology, Technical Report, February 2002.

[42] William S Noble, *"What is a support vector machine?"*, Nature Biotechnology, Volume 24, number 12, December 2006.

[43] B.E. Boser, I.M. Guyon, and V.N. Vapnik, *"A training algorithm for optimal margin classifiers."*, In D. Haussler, editor, 5th Annual ACM Workshop on COLT, 1992. ACM PRESS.

[44] Asa Ben-Hur and Jason Weston, *"A User's Guide to Support Vector Machines"*, Data Mining Techniques for the Life Sciences, Methods in Molecular Biology Volume 609, 2010.

[45] John Platt, *"Advances in Large Margin Classifiers"*, MIT Press, 1999.

Other References

[46] Department of Energy, tips on appliance use, July 5, 2012. link: energy.gov/energysaver/articles/tips-appliances

[47] The Association of Manufacturers of Domestic Appliances, Time to Change: http://www.t2c.org.uk/saving-energy/electricity-usage/