

# **A New Method for Residential Side Non-Intrusive Load Monitoring**

**Sheng Wu**

A thesis presented in fulfilment of the requirements for the degree of  
Doctor of Philosophy

Department of Electronic and Electrical Engineering  
University of Strathclyde, Glasgow, UK

## DECLARATION

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:  Date:

## ABSTRACT

This thesis proposes a new non-intrusive method for residential load monitoring. The proposed method can detect appliance switching events, separate appliance electric features, and identify appliance types. Compared with other non-intrusive monitoring methods, the proposed method improves the monitoring accuracy and decreases the monitoring response time.

Firstly, the monitoring hardware was designed and constructed to sample and acquire the aggregated electric data of one residential area.

Secondly, the sampled data were processed and analysed with the proposed method, which achieves the monitoring of individual appliance running conditions and power consumption in this area in a non-intrusive way. The data analysis process includes three steps, 1) the appliance switching event is detected by the Heuristic detection method. 2) the working current of the switched appliance is separated according to the difference method, 3) the type of switched appliance is identified with the K-nearest neighbour method according to the appliance's current harmonic components, and the identification result is modified and corrected according to appliance operation pattern with the aid of a Back Propagation Neural Network.

Thirdly, the proposed NILM method was tested through offline and online applications. The offline application involves three days of pre-recorded data which were processed and analysed. The online application consists of two parts. The first part is a direct application for four domestic homes during one day (24 hours). As for the second part, the proposed monitoring method was applied to one domestic home for ninety days. All the online and offline tests, the running conditions and the power consumption of appliances were monitored and recorded.

Due to the test results, the proposed method is reliable and offers a powerful monitoring method for demand side management.

## **ACKNOWLEDGMENTS**

I would like to express my sincere appreciation to my supervisor, Professor K. L. Lo, for his supervision and support throughout my research work. The completion of this thesis would not have been possible without his constant encouragement and patient guidance.

My gratitude also goes to my colleagues, who are and were studying inside and outside the PSRG. I am thankful for the friendship, the company and the unforgettable memories over these challenging years. And thanks to all my colleagues in NCEPU, especially Dian Jiao and Yucheng Gao.

Finally, I would like to thank my parents and sister for their strong encouragement and love throughout my Ph.D. program. I can't complete my Ph.D. project without their support.

## LIST of ABBREVIATION

AC	Air conditioning
ADL	Active deep learning
ANN	Artificial neural network
BPNN	Back propagation neuron network
CNN	Convolutional neural network
CUSUM	Cumulative sum charts
CWT	Continuous wavelet transforms
DAE	Denosing auto-encoder
DBSCAN	Density-based spatial clustering of applications with noise
DL	Deep learning
DSM	Demand side management
DWT	Discrete wavelet transform
EK	Electric kettle
FERC	Federal energy regulatory commission
FFT	Fast Fourier transform
GAN	Generative adversarial network
GE	Geyser
GSP	Graph signal processing
HMM	Hidden Markov model
KNN	K-nearest neighbour
LAP	Laptop
LIR	Log likelihood ratio
LSTM	Long short-term memory
MO	Microwave oven
NILM	Non-intrusive load monitoring
RE	Refrigerator
SBO	Surrogate-based optimization
SNR	Signal to noise ratio
SVM	Support vector machines
THD	Total harmonic distortion
VC	Vacuum cleaner

## LIST of NOMENCLATURE

### Chapter 2

$a_m(t)$	The running state of the appliance $m$ at time $t$
$P_m(t)$	The power of appliance $m$ at time $t$
$X_1, X_2, X_3, \dots, X_n$	1st to $n$ th electric data in a discrete sampled sequence
$X_{max}, X_{min}$	The max and min data in sampled sequence
$C_i$	The cumulative sum value of $i^{th}$ sampled data
$ds(i)$	The log likelihood ratio of $i^{th}$ sampled data;
$D(k, y_i)$	The difference between unknown sample $k$ and training sample $y_i$
$x_i, x_j$	Any two elements in graph of GSP
$v_i, v_j$	Any two nodes in graph of GSP
$A_{i,j}$	The relationship between node $v_i$ and $v_j$ of GSP
$S_i$	The $i^{th}$ classification label in GSP
$y_i$	The $i^{th}$ training sample of KNN model
$N_i$	The number label of $i^{th}$ training sample in KNN model
$F_1, F_2, \dots, F_L$	Each value of feature elements in one training sample
$k$	The unknown sample of KNN model
$M_l$	The number label of $l^{th}$ unknown sample in KNN model
$G_1, G_2, \dots, G_L$	Each value of feature elements in one unknown sample

### Chapter 3

$I_{rms}$	The root mean square of current
$i(k)$	The value of $k$ th sampled current data
$I_p$	The peak current value in current sequence
$f$	The utility frequency
$T$	One period
$I_{rms}$	The root mean square value of current
$I_p$	The peak current of an appliance working current
$p(k)$	The instantaneous power
$P_p$	Power amplitude of instantaneous power
$S_{pw}$	Doublet area of instantaneous power
$a_k$	The amplitude of current harmonic component
$\theta_k$	The phase angle of current harmonic component

### Chapter 4

$I_k(t)$	The working current of $k$ th appliance
$I_n(t)$	The noise current of power system

$I_b(t)$	The house current before the switching event
$I_a(t)$	The house current after the switching event
$i_{rms}$	The house current rms value in one observation period
$\Delta i_{rms}$	The difference between house current rms values
$V_1, V_2, V_3, \dots, V_L$	1st to Lth voltage data in a discrete sampled sequence
$I_1, I_2, I_3, \dots, I_L$	1st to Lth current data in a discrete sampled sequence

## Chapter 5

$\psi_m$	The training sample (the known current) in KNN model
$x_\zeta$	The unknown sample (the separated current) in KNN model
$D(x_\zeta, \psi_m)$	The Euclidean distance between unknown and training sample
$P_{x_\zeta, \psi_m}$	The Euclidean distance after normalized
$\Delta i_{rms}$	The rms value difference between two period
$I_1, I_2, I_3, \dots, I_L$	1st to Lth current data in a discrete sampled sequence
$t_l, h_j, O_s$	The inputs for input, hidden and output layer of BPNN
$Y_l, Y_{h(j)}, Y_{O(s)}$	The outputs for input, hidden and output layer of BPNN
$u_{ij}, v_{js}$	The connection weighting of BPNN neuron
$\Delta u_{ij}, \Delta v_{js}$	The update value $u_{ij}$ and $v_{ij}$

## LIST of FIGURES

Fig.1. 1 The electricity consumption from 1980 to 2021 [1] .....	1
Fig.1. 2 Schematic diagram of intrusive monitoring.....	3
Fig.1. 3 Schematic diagram of non-intrusive monitoring .....	4
Fig.2. 1 The general monitoring process .....	13
Fig.2. 2 the sampled current sequence with 1000 sample points .....	16
Fig.2. 3 Heuristic detection process .....	17
Fig.2. 4 A GSP example with four nodes .....	22
Fig.2. 5 Basic principle of KNN .....	23
Fig.3. 1 Power features comparing between different appliances .....	34
Fig.3. 2 Appliances current features comparing.....	35
Fig.3. 3 Appliances harmonic features comparing.....	36
Fig.3. 4 Schematic diagram of data acquisition device.....	38
Fig.3. 5 the steady state working currents of appliances.....	39
Fig.3. 6 Current harmonics of domestic appliance load.....	41
Fig.3. 7 Instantaneous power curve of typical load .....	43
Fig.3. 8 Appliances running period .....	45
Fig.4. 1 Appliance's connection diagram .....	49
Fig.4. 2 One hour's house current sequence .....	52
Fig.4. 3 The $\Delta$ irms(j) values compare with the threshold of 0.1 .....	53
Fig.4. 4 One hour's current sequence with low SNR.....	54
Fig.4. 5 The switching detection result under 0.1 and 1 .....	55
Fig.4. 6 The switching detection result under threshold 0.1 and additional $\Delta$ irms comparing.....	56
Fig.4. 7 Flow chart of switching event detection .....	58
Fig.4. 8 The $-90^\circ$ difference between current and voltage sequence phase angle .....	59
Fig.4. 9 The $0^\circ$ difference between current and voltage sequence phase angle...	59
Fig.4. 10 The $90^\circ$ difference between current and voltage sequence phase angle	60
Fig.4. 11 Appliance working current separating process .....	62
Fig.4. 12 The one hour's current sequence .....	63
Fig.4. 13 The switching event detecting result .....	63
Fig.4. 14 Working current sequence before and after 11 switching events.....	65
Fig.4. 15 The current separation after 1 <sup>st</sup> switching event .....	66
Fig.4. 16 The current separation after 2 <sup>nd</sup> switching event.....	67
Fig.4. 17 The current separation after 7 <sup>th</sup> switching event.....	67
Fig.4. 18 The separated current sequences after switching event .....	68
Fig.5. 1 Four appliances switching times in one week .....	72
Fig.5. 2 Four appliances switching times in ten weeks.....	74
Fig.5. 3 Separated current working wave.....	79

Fig.5. 4 The comparing of identification accuracy and period between two methods.....	82
Fig.5. 5 BPNN structure.....	84
Fig.5. 6 BPNN structure for operation pattern training .....	88
Fig.5. 7 BNPP error trend during training process.....	89
Fig.5. 8 The operation pattern curves sets.....	90
Fig.5. 9 The general operation pattern curve of a geyser.....	91
Fig.5. 10 The general operation pattern curve of a TV .....	92
Fig.5. 11 The general operation pattern curve of a laptop .....	92
Fig.5. 12 The general operation pattern curve of a microwave oven.....	93
Fig.5. 13 Geyser regression operation pattern curve and real statistical times ...	94
Fig.5. 14 TV regression operation pattern curve and real statistical times .....	95
Fig.5. 15 Laptop regression operation pattern curve and real statistical times ...	95
Fig.5. 16 Microwave oven regression operation pattern curve and real statistical times.....	96
Fig.5. 17 The flow chart of the whole identification process.....	97
Fig.5. 18 Switched-on probability of the geyser in 803rd minute .....	98
Fig.5. 19 Switched-on probability of the kettle in 803rd minute .....	98
Fig.5. 20 Switched-on probability of the laptop at the 461st minute.....	99
Fig.5. 21 Switched-on probability of the TV at 461st minute.....	99
Fig.6. 1 Schematic diagram of device and system .....	102
Fig.6. 2 Monitoring environment of office .....	104
Fig.6. 3 Total current of office from 1st hour to 8th hour .....	105
Fig.6. 4 Total current of office from 9th hour to 16th hour.....	106
Fig.6. 5 Total current of office from 17th hour to 24th hour.....	107
Fig.6. 6 The current waveform and current change point .....	108
Fig.6. 7 Separated current after laptop-B switching .....	109
Fig.6. 8 Separated current after air-conditioning A switching .....	109
Fig.6. 9 Separated current after Laptop-C, Laptop-A and Laptop-D switching	110
Fig.6. 10 Separated current after air-conditioning A state change .....	111
Fig.6. 11 Separated current after air-conditioning A state change .....	111
Fig.6. 12 Separated current after Laptop-B, Laptop-C and Laptop-D working state change .....	112
Fig.6. 13 Proportion of monitored electricity consumption in each day.....	122
Fig.6. 14 The 1st hour to 8th hour total current waveform of home A .....	124
Fig.6. 15 The 12th hour to 16th hour total current waveform of home A.....	125
Fig.6. 16 The 17th hour to 24th hour total current waveform of home A.....	126
Fig.6. 17 Appliance power consumption proportion.....	133

## LIST of TABLES

Table 3. 1 Notation and description of extracted features.....	32
Table 3. 2 Feature values of steady state current for each appliance .....	40
Table 3. 3 THD value for each appliance.....	42
Table 3. 4 Instantaneous power feature of each appliance.....	44
Table 4. 1 Switching event detecting result .....	52
Table 4. 2 The real switching event point and the detected switching point.....	57
Table 4. 3 The real switching point and the detected switching point .....	64
Table 5. 1 The UK-DALE data set details [21].....	71
Table 5. 2 First to fourth harmonics of eight appliances .....	76
Table 5. 3 The Euclidean distance between 8 training samples to each other ....	77
Table 5. 4 The switching time point of 9 unknown appliance .....	79
Table 5. 5 The frequency feature aggregation of separated currents .....	80
Table 5. 6 Euclidean distance between the training sample and separated current .....	80
Table 5. 7 The normalized Euclidean distances the training sample and separated current.....	81
Table 5. 8 Identification result based on KNN method.....	81
Table 5. 9 The number of switching times of four appliance [UK] .....	86
Table 5. 10 The number of switching times of four appliances [UK].....	93
Table 5. 11 The nine separated current identification result .....	100
Table 6. 1 Appliance name and type .....	104
Table 6. 2 The appliance switching point and time point .....	113
Table 6. 3 The Switching event classifying.....	114
Table 6. 4 The identification result of 229 switching events(3days) .....	115
Table 6. 5 The monitored and real switching condition.....	120
Table 6. 6 The monitored power consumption.....	121
Table 6. 7 Appliance name and type .....	123
Table 6. 8 The appliance switching event and switching time point of Home A (24hours).....	127
Table 6. 9 The identification result of 128 switching events in Home A.....	128
Table 6. 10 The monitored and real switching condition in Home A .....	131
Table 6. 11 Power consumption of appliances in four homes.....	131
Table 6. 12 Appliance power consumption .....	132

# CONTENT

LIST of ABBREVIATION .....	i
LIST of NOMENCLATURE.....	ii
LIST of FIGURES .....	iv
LIST of TABLES.....	vi
Chapter 1 Introduction .....	1
1.1 Research Background.....	1
1.2 Residential Load Monitoring Methods.....	2
1.2.1 Intrusive load monitoring way.....	3
1.2.2 Non-intrusive load monitoring (NILM) .....	3
1.3 Motivation .....	5
1.4 Aims and Research Objectives.....	7
1.5 Original Contribution of the Thesis.....	8
1.6 Thesis Organization.....	10
1.7 Publication.....	11
Chapter 2 Literature Review and Methodology of Non-intrusive Load Monitoring .....	12
2.1 Introduction .....	12
2.2 Overview of Non-intrusive Load Monitoring Research.....	12
2.3 Methodology of Monitoring Household Appliances with a Non-intrusive System	14
2.3.1 Data Acquisition .....	14
2.3.2 Switching Event Detection.....	16
2.3.3 Appliance Feature Separation and Extraction .....	20
2.3.4 Appliance Type Identification.....	21
2.4 Proposed Monitoring Approach .....	28
2.4.1 Switching Event Detecting with Heuristic Method.....	28
2.4.2 Separating Current Feature with Difference Method.....	29
2.4.3 Accurate Identification of Appliances with Similar Feature .....	29
2.5 Summary .....	30
Chapter 3 Appliance Features Analysis and Extraction for NILM.....	31
3.1 Introduction .....	31
3.2 Researching of Household Appliance Electric Feature.....	31
3.3 Appliance Electric Feature Extraction.....	37

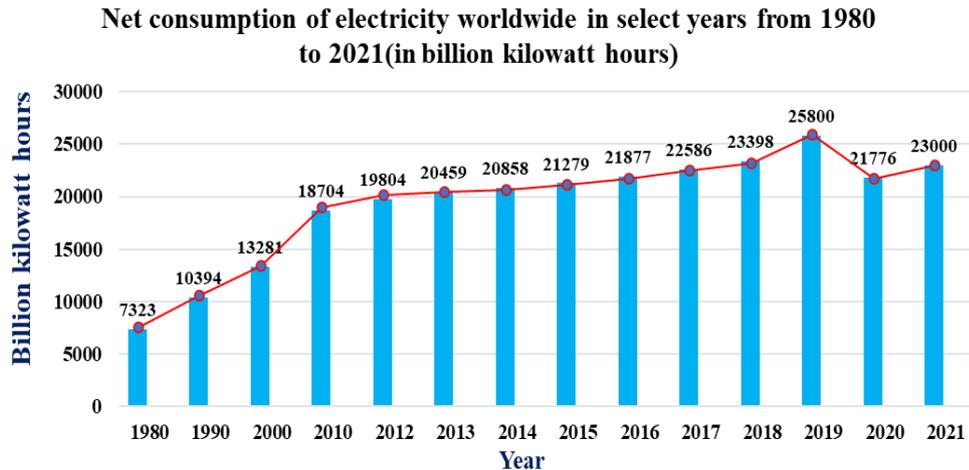
3.3.1	<i>Appliance Steady State Working Current</i> .....	38
3.3.2	<i>Appliance Working Current Harmonics</i> .....	40
3.3.3	<i>Appliance Transient and Steady Power Features</i> .....	42
3.4	Non-electric Feature Introduction .....	44
3.5	Summary .....	46
Chapter 4	Appliance Switching Event Detection and Working Current separation .....	48
4.1	Introduction .....	48
4.2	The Character of Running Household Appliances .....	48
4.3	Appliance Switching Event Detecting Based on Heuristic Detection Method .....	50
4.4	Separating the Appliance Working Current through the Difference Method .....	58
4.5	Testing of Switching Event Detection and Working Current Separation .....	63
4.6	Summary .....	68
Chapter 5	The Switched Appliance Type Identification .....	70
5.1	Introduction .....	70
5.2	Analyse the usage habits of customer and the appliance running pattern .....	71
5.3	Identification of Appliance Type .....	75
5.3.1	<i>The KNN Method reconstructing and separated current identifying</i> .....	75
5.3.2	<i>Identification Testing Based on KNN</i> .....	78
5.4	Modification of Identified Result Based on Appliance Operation Pattern .....	83
5.4.1	<i>BPNN Principle for Operation Pattern Regression</i> .....	84
5.4.2	<i>Operation Pattern Regression Based on BPNN</i> .....	85
5.4.3	<i>Verification and Testing of Operation Pattern Curve</i> .....	93
5.4.4	<i>Identification Correction Testing</i> .....	96
5.5	Summary .....	100
Chapter 6	Application of Proposed Method .....	102
6.1	Introduction .....	102
6.2	Monitoring Hardware .....	102
6.3	Off-line Application Based on 3 days' recorded Data .....	103
6.3.1	<i>Appliance switching detection and type identification</i> .....	105
6.3.2	<i>Power consumption monitoring</i> .....	121
6.4	One Day On-line Test (24hours) .....	122
6.4.1	<i>Appliance switching detection and type identification</i> .....	123

6.4.2. Power consumption monitoring.....	131
6.5 Three Months Monitoring of Power Consumption for a Normal Home .....	132
6.4 Summary .....	133
Chapter 7 Conclusion and Future Work .....	135
7.1 General conclusions .....	135
7.2 Suggestions for future research .....	136
Appendix I Data: Office current waveform from 25th hour to 72nd hour.....	138
Appendix II Data: Home B, Home C and Home D current waveform.....	144
Appendix III: The identification result of appliance switching in Home B, Home C and and Home D.....	153
Reference .....	161

# Chapter 1 Introduction

## 1.1 Research Background

Global electricity consumption has been increasing steadily in domestic, commercial, and industrial sectors around the world. Fig.1.1 shows the consumption of electricity worldwide, ranging from 7323 billion kWh in 1980 to 23000 billion kWh in 2021. The general trend of consumption is indicated by the red line which shows a steady increase except for the year 2020 and 2021 where it shows a slight dip when compared to all previous years. The slight dip can be accounted for by the effect of Covid-19. Both personal and commercial activities in many countries were partially interrupted during those years resulting in a dip in energy demand.



**Fig.1. 1 The electricity consumption from 1980 to 2021 [1]**

The continuous increase may lead to a shortage of electric power [2], therefore it is essential to bridge any demand and supply gap as the trend requires [3]. One way is to expand the generation, transmission, and distribution system capacity [4]. However, there are difficulties in building new power stations and new transmission lines due to: (1) the availability of suitable sites, and (2) the attainability of access permission providing ‘way leave’ for transmission lines. All restrictions are a result of environmental protection policies [5, 6]. The concept of demand side management (DSM) is suggested as a way to alleviate power shortage and bridge the demand supply gap since it encourages customers to amend decisions about their energy usage pattern

and enables energy suppliers to minimize the peak demand overload and reshape the load profile [7, 8]. The benefit of DSM has been reflected in industrial sectors [9] in countries such as China, where it has been applied in the building industry since 2004. The total electricity consumption within China's building industry was 590 billion kWh, with the construction of large public buildings accounting for 30% of the electricity consumption in that industry. The implementation of effective DSM in large public building construction can effectively save 5% - 10% of electricity from being wasted and also improve power transformation and distribution efficiency[10].

When DSM is applied in the residential sector, although the total demand is comparable to the industrial sector, the number of residential loads is large, the power of each is low, which means that it is meaningless to manage a single residential load through the same type of DSM application as one would use in an industrial sector. To improve DSM efficiency and feasibility in the residential sector, obtaining individual load operational characteristics and power consumption totals is helpful since the aggregated control of load cluster could be an effective residential demand management tool [11, 12], a tool which would require accurate monitoring of a single residential load. However, in a residential home, the number of electrical appliances can easily be over a dozen, and the operation of these appliances is random. Several appliances are always operating in parallel, so the traditional monitoring way is unable to differentiate the power consumption of individual appliance and track appliance run times [13, 14]. This tracking only monitors the total electricity consumption of the customer within a period through energy meters [15]. To obtain the data needed for residential DSM, new residential load monitoring methods have been previously proposed.

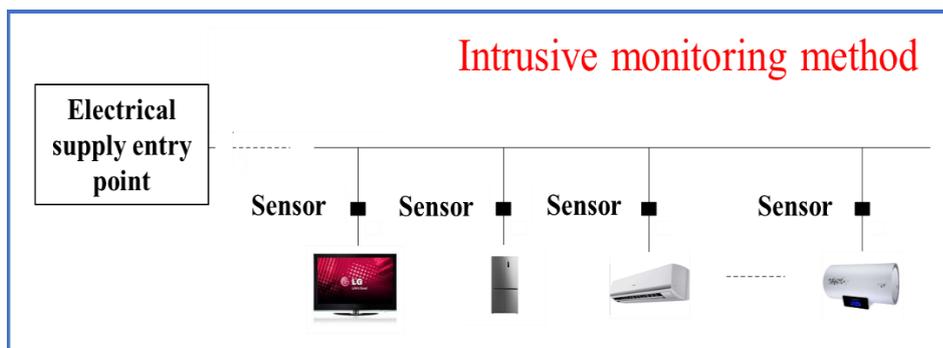
## **1.2 Residential Load Monitoring Methods**

In residential load monitoring, two main methods are direct and indirect monitoring. The direct approach involves the installation of sensors directly connected to the appliance, the sensors are embedding in each appliance, so it is an intrusive way of load monitoring. The indirect approach analyses the current, power and other electric parameters flow going into a residential home and then deduces what appliances have operated and their individual energy consumption[16, 17]. The latter

is a computational technique based purely on analytical tools, so it is a non-intrusive way of load monitoring [18, 19].

### ***1.2.1 Intrusive load monitoring way***

Intrusive monitoring installs sensors for each appliance being monitored. The appliance operation status and power consumption information are obtained by directly measuring and recording the usage. Fig.1.2 shows the schematic diagram of intrusive monitoring.



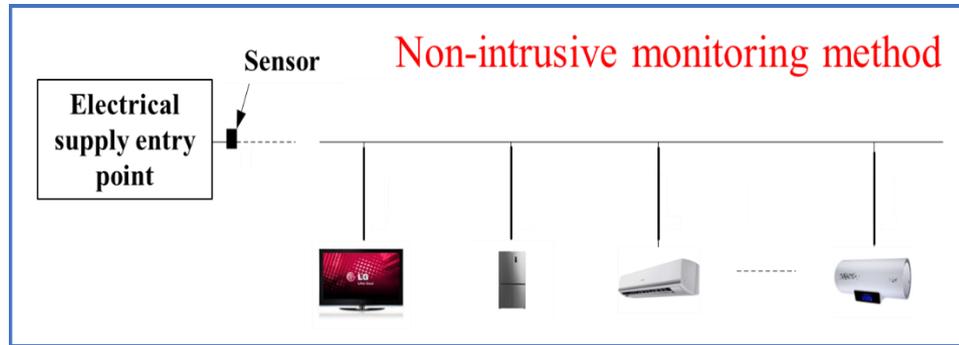
**Fig.1. 2 Schematic diagram of intrusive monitoring**

Due to the different data needed, multiple dedicated sensors are used, such as Drenker [20] uses power sensors to monitor household appliances that are running and provides power measurements, and Kelly [21] uses sensor sets to monitor the working current, phase angle, active and reactive power of the specific appliance. Although the intrusive monitoring system has high accuracy and reliability, the implementation cost is high, and it is challenging to establish the sensor-distributed measurement network in actual practice [22]. Further, the installation and expansion of sensors cause inconvenience for customers. Intrusive monitoring is not a feasible solution for widespread residential load monitoring, it is more suitable for specific high-power consumption appliances.

### ***1.2.2 Non-intrusive load monitoring (NILM)***

All household appliances in residences are connected in parallel at the electricity supply point, which means switching ON and OFF these appliances changes the

current, power, and other electric parameters [23] at the electricity supply point. NILM can monitor household appliances' operating conditions and power consumption by analysing these electric parameters. Compared to intrusive monitoring which uses multiple sensors, NILM uses a single sensor. Fig.1.3 shows the schematic diagram of a household appliance monitored using NILM.



**Fig.1. 3 Schematic diagram of non-intrusive monitoring**

The technique of NILM was first applied to monitor residential appliances by Hart [24] in the 1980s. Data of the total active and reactive power (P-Q) was measured and collected at the electrical supply point of a home. The P-Q changes were then analysed and mapped into a power feature space to determine which appliance was responsible for the change. Since then, the technique has evolved to include more electric parameters and complex data analysis processing to improve monitoring accuracy [25].

Reference [26] showed how weighted images generated by different Voltage-Current (V-I) curves were used as input for a convolutional neural network (CNN) to identify the working current of various appliances. The accuracy obtained for the public datasets (PLAID and WHITED) was around 80%. References [27] and [28] used Active Deep Learning (ADL) and Discrete Wavelet Transform (DWT) to extract and identify appliance features from the current data at the electrical supply point, achieving an accuracy of above 90% through testing three public datasets. In reference [29], a combination of Generative Adversarial Network (GAN) and CNN was used to denoise the current data and extract the appliance feature at the electrical supply point, which resulted in a monitoring accuracy of around 92%. Reference [30] utilized a deep neural network that combined a regression sub-network with a classification sub-

network based on the power data and status information of the appliance to disaggregate the mixed power data at the electrical supply point. Finally, reference [31] proposed a brute-force method and a greedy method to disaggregate the mixed power data at the electrical supply point into the individual appliance level.

Although monitoring accuracy has improved, there remains three limitations in the NILM application. Based on references [25]-[31], the first limitation is that monitoring takes a long time due to complex computation and data processing. The efficiency of identifying and monitoring the appliance is reduced. When frequent switching events occur, another appliance may also be activated during the data processing period and affect the final monitoring results. The second limitation is related to the design of the feature extractor, classifier, and decomposer. The features used to train and construct corresponding models are empirically selected, which can lead to fault monitoring if inappropriate features are used in the training and constructing process. The third limitation is low monitoring accuracy for appliances with the similar electric features.

### **1.3 Motivation**

Power shortage has become a serious problem in the current power system, which cannot be resolved solely by enlarging the power supply. Demand-side management (DSM) utilizes monitoring and control technologies to innovate traditional power systems and manage power distribution, making it a potential solution to tackle this challenge. DSM can achieve several goals, such as providing customers with detailed energy efficiency information, illustrating the impact of customer behaviour on energy efficiency, and guiding them to take corresponding energy-saving measures. For power suppliers, DSM can help improve smart grid implementation, relieve power supply shortages, ensure power grid stability, balance power supply and demand, and analyse the composition of electric energy consumption, thereby providing reliably sourced power.

Therefore, more and more policies and regulations encourage DSM. The EU energy efficiency directive allows everyday residential consumers to participate in power management, either alone or by aggregating scattered loads [32]. Furthermore, in the UK, Scottish Power cooperates with Honeywell (using their developed

automatic demand responses for industrial and commercial equipment) to improve power efficiency [33]. In the United States, with the Energy Policy Act [34] and the American Clean Energy Security Act [35] being promulgated, several major power companies and market institutions have begun to implement demand-side power management. In 2013, power transmission network operators could choose the auxiliary service directly from a third party and the settlement mechanism, including load management [36]. In 2018, the Federal energy regulatory commission (FERC) issued Act No. 841, which provided that the corresponding power management service can participate in the power market through market competition [37]. China, in 2015, specified detailed rules for demand-side power management concerning the background, principles and objectives of power management implementation [38]. In 2016, the plan for electric industry development from 2016 to 2020 emphasized the significance of power management [39]. In 2017, the government issued new rules for power demand-side management. The practice of DSM operation is beneficial in environmental protection, green power, intelligent electricity, and promoting renewable energy enriched by the development of DSM. In 2020, the government and State Grid company issued the power reliability management rules to strengthen power management regulations, further giving DSM a higher focus on power system reliability [40].

As rules encouraging DSM development become more prevalent and researchers focus more on DSM, it will eventually become widely applied in power systems [41]. However, any load management system must be based on a correct understanding of individual appliances' running characteristics and power consumption conditions, making load monitoring technology increasingly vital in DSM. NILM is thus gaining more attention for two reasons: 1) NILM is more suitable for widespread deployment since the non-intrusive approach only requires the current, power, and other electric data to be analysed at the electrical supply entry point; 2) the power consumption and running condition of each appliance within the monitoring area can be obtained more economically. To widely apply NILM in the residential load monitoring, its weaknesses in accuracy, practicability and efficiency must be overcome. Here are some of the common weaknesses of NILM and the proposals on how to overcome them:

- Faulty monitoring results are always caused by using inappropriate appliance features to train and construct the NILM model. To improve the NILM accuracy, effective appliance features must be selected and analysed before the model and method design, to filter out invalid features for load identification in advance.
- The computation process of NILM is too complex, resulting in a longer response time for monitoring. Another appliance may start running and affect the monitoring and results. This concern is negligible under the intrusive method. To reduce the response time of NILM, minimize the analysis of duplicate data, and simplify the calculation process of existing methods.
- The accuracy of non-intrusive monitoring is significantly challenged, especially in the monitoring of appliances with similar electric features. To improve monitoring accuracy, a combination of electric and non-electric appliances features can be used in the monitoring process.
- Most NILM models are tested using public datasets, which are not actual data in residential power environments, and thus are insufficient and unconvincing. To increase the credibility of testing, public datasets are combined with real data to test NILM models.

#### **1.4 Aims and Research Objectives**

This thesis investigates the potential of using a non-intrusive method for monitoring electric consumption in residences to replace the intrusive monitoring system. The aim is to improve the monitoring accuracy, decrease the monitoring response time, and enhance the practicability of NILM when applied in residential settings. Here are four main objectives:

- The first objective is to offer an updated review of the appliance features for load identification, as previous research has often randomly or empirically selected electric features for model construction and training. A systematic analysis and comparison of comprehensive features will ensure that the selected features are valid in NILM, which will help optimize later methods and model designing, making the monitoring process more accurate.

- The second objective is to reduce the computational complexity and the response time of NILM. The disparate and variational data which occurs in switching on and off events are more valuable than reduplicative and invariable data. By separating and identifying only the data around the switching events, the computation process of NILM will be much simplified, and the NILM response time is reduced. Furthermore, the separating and identifying methods will redesign and improve to further reduce NILM response time.
- The third objective is to improve monitoring accuracy. As electric features are similar in many different appliances, no matter how much the model tries to optimize and train the identification process, the overlapping of electric features will inevitably cause faulty identification and decrease monitoring accuracy. Combining the appliance operation pattern into the monitoring can overcome this issue and largely improve the monitoring accuracy.
- The fourth objective is to offer a comprehensive testing method to check the practical application of NILM. Although there are many public datasets that can be used to test monitoring methods, the entire monitoring process is done in a computer program, which idealizes or downplay constraints and problems in actual practical applications. Testing the proposed NILM method using power data from an actual residential home and the power data from public datasets, which will provide more genuine results and allow a better understanding of the practical application of NILM.

### **1.5 Original Contribution of the Thesis**

This thesis makes several original contributions which are summarized below with techniques used to achieve the research objectives:

- Original Contribution 1: Chapter 3 conducts a systematic comparison and analysis of appliance features to select the most effective features for NILM. By extracting and analysing features from frequently used household appliances, the most effective ones are combined and selected. The selected features in this thesis are actual measuring, which is used for the further monitoring process. Unlike other

researchers who did not carry out appliance feature analysis, this thesis places equal importance on both appliance features analysis and model establishment. This reduces the faulty monitoring results caused by the inappropriate selection of features.

- Original Contribution 2: Chapters 4 and 5 propose a monitoring method with a short response time and reduced computation complexity. Firstly, in Chapter 4, the Heuristic method is improved, which is used for fast detection of appliance switching events. The separation method of the current data being before and after the switching event is designed and improved to avoid the unnecessary analysis of reduplicative and invariable data. Secondly, in Chapter 5, the KNN model is reconstructed and trained based on the prior appliance feature analysis in Chapter 3. This reduces the storage space required for the model and simplifies its structure as well as the data process program.
- Original Contribution 3: Chapter 5 proposes a correction method for monitoring results to consider appliance operation features in NILM. This overcomes the problem of appliance electric features overlapping in the monitoring process. A BPNN is trained with appliance non-electric features to fit the appliance operation pattern, which is used as a reference for correction purposes. Compared to other research, this correction strategy improves monitoring accuracy, and the BPNN is only used as an offline technology to obtain the appliance operation pattern. This approach incorporates the better regression ability of BPNN, without increasing the complexity of the monitoring method.
- Original Contribution 4: Chapter 6 tests the proposed monitoring process with both offline data and online data measured in an actual working environment. The offline test involves 3 days of pre-recorded data sampled directly from a real residential environment which proved the feasibility of the proposed method. The online test consists of two parts, which include a direct application to 4 domestic homes for a day (24 hours) to test the robustness of the proposed method, and another direct application to 1 domestic home for 90 days to test the reliability of the proposed method. Compared with testing of published papers on monitoring methods that use public datasets and where the whole monitoring process is

completed in the computer program, this testing method as proposed in this thesis to be carried out under a real and practical environment is more comprehensive and can effectively check the advantages and disadvantages of the NILM.

## **1.6 Thesis Organization**

This thesis focuses on NILM based on the working current of a single appliance. A complete framework for this non-intrusive load monitoring system is designed using high-frequency sampling. The framework includes research on the separation and identification of working currents of appliances and the correction of load identification results to improve monitoring accuracy. The proposed algorithm is tested under practical environments to demonstrate its effectiveness. The thesis is structured as follows:

Chapter 2 presents a general overview of NILM, including various solutions for data acquisition and analysis. The methodology of each solution is introduced, and their advantages and disadvantages are compared. Finally, the proposed method is presented.

Chapter 3 analyses and compares various appliances' features, which is the prior research for NILM, the features with higher identification are selected as the main contents for further research. The selected features are measured and extracted from the actual residential environments.

Chapter 4 develops fast switching events detection and features separation methods. Firstly, the Heuristic method is improved to achieve fast switching event detection; when the switching event is detected, the working current of a switched appliance is separated through the modified difference method.

Chapter 5 presents a method to improve the identification accuracy of loads with similar electrical characteristics. Firstly, based on the frequency components of different appliance working currents, the separated currents are transferred from the time domain to the frequency domain to identify the separated currents using the KNN method. The KNN model is reconstructed and trained. Secondly, home appliances have different switching probabilities in a day. A BPNN is trained, based on the running period and times of known appliances, to fit the operation pattern curve of these appliances. The operation pattern curve reflects the switching probability

distribution of the appliance, and then the identification results are modified based on the switching probability.

Chapter 6 is the application of the proposed method. Different appliances of 4 houses are monitored, the power consumption of each appliance is obtained, and the running of each appliance is tracked.

Chapter 7 draws conclusions and provides recommendations for future work.

## **1.7 Publication**

Sheng Wu and K.L.Lo, "Non-intrusive monitoring algorithm for resident loads with similar electrical characteristic," in *Processes*, Vol. 08, Iss. 11, Art No. 1385, DOI: 10.3390/pr8111385, 2020.

Sheng Wu and Liya Liu, "Research on Features of Residential Loads and Establishment of Feature Library," in 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), DOI: 10.1109/EI2.2018.8582382.

## Chapter 2 Literature Review and Methodology of Non-intrusive Load Monitoring

### 2.1 Introduction

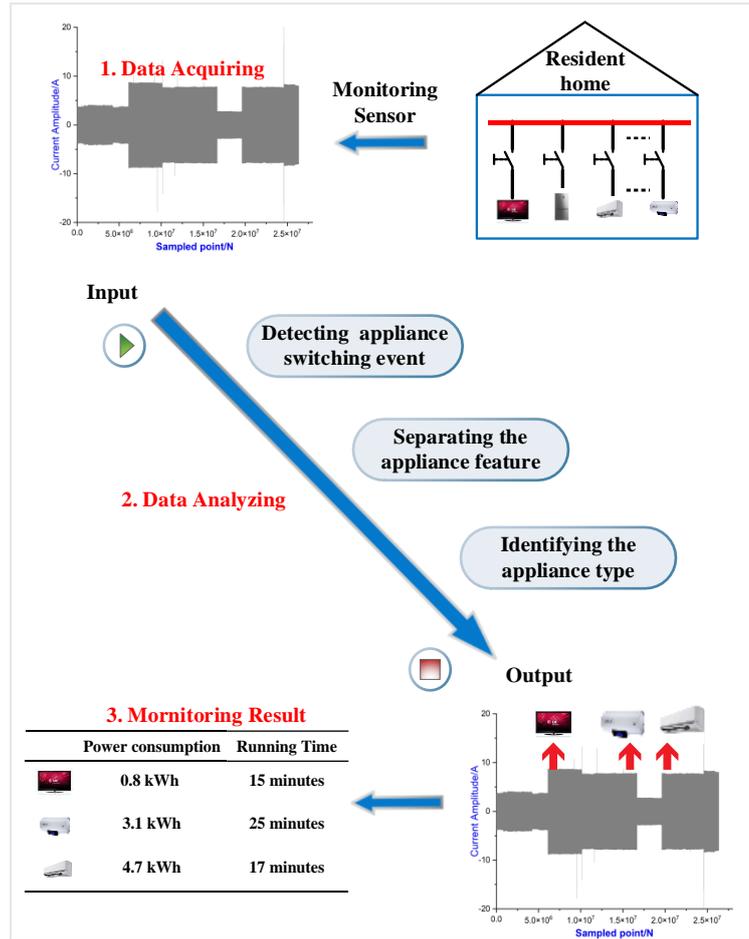
This chapter conducts a literature review on the methodology of non-intrusive monitoring for residential appliances. First, the process of NILM is explained in Section 2.2 which involved two tasks: to acquire the integrated and accurate electric data at the power supply entry point; to analyse the obtained data for determining the appliances running states and power consumption characteristics. In Section 2.3, two common data acquisition methods for NILM are introduced: high-frequency sample and low-frequency sample. In Section 2.4, the data analysing process is introduced, which consists of the following three steps: 1) appliance switching detection, 2) switched appliance feature separation, and 3) appliance type identification. The advantages and disadvantages of the different solutions for each step are then compared. In Section 2.5, the proposed monitoring model is introduced.

### 2.2 Overview of Non-intrusive Load Monitoring Research

The purpose of NILM is to obtain running states and power consumptions of individual appliances from the data collected at the electricity supply point. The total power (the power of the electricity supply point) at any specific time is determined by the running states and corresponding power of each appliance at that time. The total power  $P(t)$  at time  $t$  can be expressed as:

$$P(t) = \sum_{m=1}^M a_m(t) * P_m(t) \quad (2.1)$$

$M$  is the number of running appliances at time  $t$ .  $a_m(t)$  represents the state of the appliance  $m$  at time  $t$ ,  $a_m(t) = 1$  (the appliance is running steadily at time  $t$ ), otherwise,  $a_m(t) = 0$ .  $P_m(t)$  is the power of appliance  $m$  at time  $t$ . Therefore, by processing the obtained data at the power supply entry point and determining the type of appliance, the  $a_m$  and  $P_m$  at each time  $t$ , the running states and power consumptions of every appliances can be monitored. Fig.2.1 illustrates the overall monitoring process:



**Fig.2. 1 The general monitoring process**

As shown in Fig.2.1, the electric data is first acquired at the electricity supply point by the monitoring sensor. Next, this acquired data is processed through the detecting, separating and identifying steps to obtain the type of appliance, the power  $P_m$  for each appliance and their appliance states  $a_m$ . Finally, based on the type of appliance, the  $P_m$  and  $a_m$  values, the power consumption and running period of each appliance are monitored.

This data acquisition is essential to ensure the authenticity of the electric parameters collected, and the acquired data should contain a sufficient number of appliance features. Without accurate and integral electric data, the subsequent data analysis becomes meaningless due to data distortion and insufficient appliance features.

For data analysis, as shown in Fig.2.1, the switching detection is the start of the monitoring, which affects the sensitivity of monitoring. The detection result also

affects the value of  $a_m$ , if an appliance was switched and the switching event was detected, this appliance state  $a_m$  will be set as 1. The separation method affects the integrity of the appliance feature, the separated features is used in the appliance type identification and power  $P_m$  calculation. The identification method affects the accuracy of monitoring. Thus, any errors in detecting, separating and identifying will affect the monitoring results (the type of appliance, the power  $P_m$  of appliance, and the running state  $a_m$  of appliance). To improve the monitoring accuracy, it is possible to use more sophisticated and advanced methods in each step. However, complex analysis processes may not be feasible for online NILM applications due to: 1) limited computing power of online NILM, 2) significant increase in the response time for monitoring. Therefore, a better monitoring model is required to combine all three aspects of detecting, separating and identifying, and better allocate the limited computation power for accurate and rapid monitoring.

Lastly, it is important to ensure the final calculated results for the power consumptions and running periods are as close to the actual values and the error margins are not too large.

### **2.3 Methodology of Monitoring Household Appliances with a Non-intrusive System**

During the monitoring process of a household's appliances, the final monitoring accuracy and the computation complexity are affected by factors such as data acquisition, appliance switching detection, switched appliance feature separation and appliance type identification. The following sections shall summarize and compare the advantages and disadvantages of these four factors.

#### ***2.3.1 Data Acquisition***

At the power supply entry point of any residential home, only the voltage, current, active power and reactive power can be directly measured and acquired. As for the other electric parameters, such as the frequency components in the power spectrum, current harmonic waves, the apparent electromagnetic interference noise, the microscopic value changes in voltage and the impact peak in the current, they must all be derived and calculated from the measured current, voltage and power data. The

corresponding parameters will inevitably be distorted in the process of extraction and calculation, which may affect the accuracy of data analysis. Hence, the current, voltage and power are the three most suitable parameters used in NILM, as their corresponding data can be acquired directly at low and high frequencies that range from 1Hz to 10kHz.

For low-frequency NILM, the sample frequency is usually in the range of 1Hz to 60Hz. This simplifies the hardware inventory, making the data transmission and storage more convenient. However, low-frequency sampling in NILM is particularly challenging due to several reasons. Firstly, the change of appliances' operating states is instantaneous, low frequency sampling may miss some data during transient states [42]. Secondly, noise from unknown appliances, abnormal transients and load fluctuations can distort the acquired data. Having few samples makes it difficult to distinguish between real data and distorted data. Thirdly, the appliance features may have serious overlapping when sampled at low frequency. Lastly, the authenticity of the electric parameters may be unguaranteed. References [43] and [44] show that although power data is acquired from one (same) appliance using different low-frequency sample devices, the obtained data is not totally identical. The collected data from each different device always has a 10-20% rate of error, which may create more errors in subsequent data analysis.

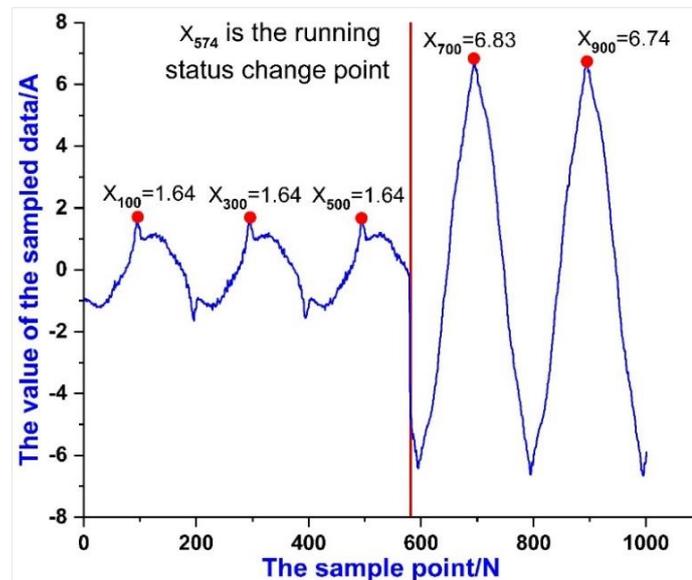
As for high-frequency NILM, the sample frequency is usually in the range of kilo-Hz or mega-Hz. The obtained data is more refined, elaborate, and accurate. Although the obtained information is abundant and various appliance features can be separated from the high-frequency sampled data, the corresponding device usually draws a high cost and requires a special design. This is raised in reference [45], where the researchers designed a low-cost data acquisition system called YoMo (You Only Measure Once). The system uses current transformers to measure the current value in a non-intrusive way. The accuracy and resolution of the YoMo system can also meet various load monitoring requirements from 4W to 5kW. In reference [46], an intelligent digital power acquisition and analysing platform called C-meter (Cognitive Power Meter) was designed to support multichannel high-frequency data acquisition. The hardware requirements and costs are relatively high for this high-frequency NILM.

Comparing the low-frequency and high-frequency sampling methods in NILM, the former method is more suitable for performing monitoring where there is less

system noise and the running of appliances is constant. The latter method is a better choice when the appliance running is more random and the system noise is high. This thesis focuses on the high-frequency sampling method.

### 2.3.2 Switching Event Detection

After the electric parameters (current, voltage or power) at the electricity supply point are sampled with high or low frequencies, the obtained data is presented in a discrete sequence  $(X_1, X_2, X_3, \dots, X_n)$ , where  $X_1$  to  $X_n$  are the corresponding values of each sampled data. As shown in Fig.2.2, the obtained current sequence includes 1000 sampled data.



**Fig.2. 2 the sampled current sequence with 1000 sample points**

As seen from Fig.2.2, the sampled data becomes reduplicative and invariable when the running status of appliances is stable, for example, the values of  $X_{100}$ ,  $X_{300}$  and  $X_{500}$  are the same. This suggests that most of the acquired data is redundant and only the data surrounding the switching event is useful for this monitoring. Therefore, knowing the appliance status change point and confirming the appliance in steady operation before disaggregating or identifying the sampled data [47] become essential in decreasing the response time and computational complexity of monitoring. The main objective of the switching event detection is hence to determine whether the

appliance running status has changed and has been steady. The detection can be achieved in two ways: firstly, by judging whether the sampled data is stable within a fixed period, and secondly, by detecting the mutation point from the sampled data.

### 2.3.2.1 Judging Data Sequence Stability

Heuristic detection [48] is the main method for judging data sequence stability in NILM. First, an observation period is determined, with one observation period including  $n$  sampled data. Next, the sampled data in the observation period forms a data set  $[X_1, X_2, X_3, \dots, X_n]$ . Finally, the difference between the maximum data  $|X|_{max}$  and minimum data  $|X|_{min}$  within one data set is used to obtain the change amount. If the change amount is less than a certain threshold  $\varepsilon$ , this observation period is considered as a stable region, which indicates that the switching event has not occurred. If the change amount is greater than the threshold, the corresponding observation period is defined as the change area, where the switching event has occurred. Fig.2.3 shows an example in which the entire detection process includes  $6 \times 10^6$  sample points, with four observation periods. The change between  $|X|_{max}$  and  $|X|_{min}$  for each observation period is 5.49, 0.2, 0.09 and 1.5 respectively. The change for Period 1 and Period 4 is significant, so they are identified as the change area where the switching event has occurred.

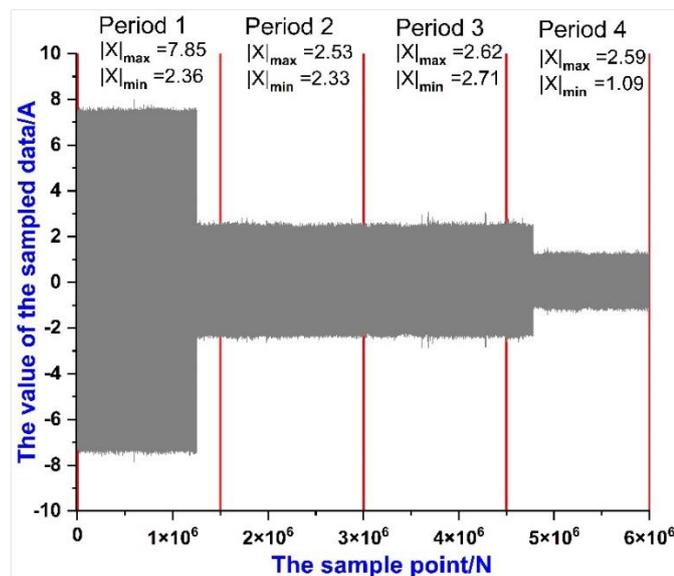


Fig.2. 3 Heuristic detection process

The principle of Heuristic detection is simple with low consumption [48], so the response time of a switching event detection is almost in real-time. However, there are two disadvantages when it is used in NILM. The first one is that the auxiliary time meter is required during the observation period. However, the coordination between the auxiliary time meter and the data acquisition device is unsatisfactory, which can lead to an increase in the complexity of the hardware and monitoring process. The second disadvantage is the limited ability to detect the switching event of small power appliances with negligible data change. The inability to detect switching events of these small power appliances decreases the sensitivity of this Heuristic detection method.

### 2.3.2.2 Detecting Mutation Data

The cumulative sum (CUSUM) [49] and log likelihood ratio (LIR) [50] are commonly used methods to detect mutation points in NILM. Instead of determining whether sampled data within an observation period is stable, these methods process each sampled data by calculation to detect if they are the mutation points. This enables a higher detecting sensitivity.

The CUSUM is a sequential analysis approach which determines how the samples vary from their mean values based on the cumulative sum value. The cumulative sum of the samples can be calculated using the following functions:

$$Mean_n = \frac{1}{n} (\sum_{i=1}^n X_i) \quad (2.2)$$

$$C_n = C_{n-1} + (X_n - Mean_n) \quad (2.3)$$

where  $X_n$  is the  $n^{th}$  sampled electric data (such as real power, reactive power and current);  $Mean_n$  is the average value of sampled data from 1<sup>st</sup> to  $n^{th}$ , and  $C_n$  is the  $n^{th}$  cumulative sum value of  $n^{th}$  sampled data. The initial cumulative sum value  $C_0$  is taken to be zero; the mutation data is obtained depending on the abrupt changes in the cumulative sum. When applying CUSUM to online residential appliance monitoring, one disadvantage is that the residential site's current, power and voltage are all sinusoidal signals. Although the appliance running state is stable, the  $Mean_n$  of

sampled data changes and the cumulative sum must be modified, leading to more errors in the detection process.

As for the LIR method [51], it relies on the change of sampled data mean value to calculate the log likelihood ratio when the mean value of electrical parameters changes beyond a specific value, it can overcome the problem of cumulative sum modifying in CUSUM. The main calculation process is as follows:

$$ds(i) = \begin{cases} \frac{\mu_{i-1} - \mu_{i+1}}{\sigma^2} \times \left| X_i - \frac{\mu_{i+1} - \mu_{i-1}}{\sigma^2} \right|, & |\mu_{i+1} - \mu_{i-1}| > P_{th} \\ 0, & |\mu_{i+1} - \mu_{i-1}| \leq P_{th} \end{cases} \quad (2.4)$$

where  $X_i$  is the  $i^{th}$  sampled data;  $ds(i)$  is the LIR of  $i^{th}$  sampled data;  $\mu_{i-1}$  and  $\mu_{i+1}$  are the mean values of sampled data sets before and after  $i^{th}$  sampled data.  $\sigma^2$  is the variance;  $P_{th}$  is the threshold. The mutation data within a data set is detected using the  $ds(i)$  value. If the log likelihood ratios in a data set are not all zero, then the data with the maximum  $|ds(i)|$  is the mutation point of this set. The disadvantage of this method is that the set of threshold value  $P_{th}$  is difficult when the system noise is high and there is a large number of appliances. When the  $P_{th}$  value is too small, the noise may be detected as a switching event. On the other hand, when the  $P_{th}$  value is large, some switching events may be missed.

While mutation data can be detected by the CUSUM and LIR calculations to indicate the occurrence of switching events, the number of such data in power systems becomes too huge. Especially since NILM is based on high-frequency sampling, every minute of the sampled data contains a mutation point. This would require the removal of irrelevant points. Intelligent algorithms are thus used to determine whether the mutation point is caused by the appliance switching or the operation state changing. Reference [52] builds on the LIR approach by combining it with the unsupervised clustering algorithm. It learns to adjust the parameters during the calculation process through immediate experience. In reference [53], an improved event detection method based on CUSUM is proposed. The concept of Surrogate-Based Optimization (SBO) is used instead to improve the robustness and efficiency of CUSUM. However, even though this helped to remove some irrelevant mutation data, the number of mutation points is still large and most of them are not relevant to the switching event, the efficiency of detection is low.

Comparing the CUSUM, LIR and Heuristic detection methods, the Heuristic detection method is the lowest in sensitivity and has the shortest response time for identifying the switching points of the appliances. If the detecting sensitivity could be improved and the observation period measured without the auxiliary time meter, the Heuristic detection method will be an effective tool for detecting appliance switching points.

### 2.3.3 Appliance Feature Separation and Extraction

Once the appliance switching events are detected, it is necessary to extract or separate a set of high-quality features from the sampled data to identify the appliance types and calculate the appliance power  $P_m$ . The extraction or separation will derive a feature or combination of features that can uniquely represent the individual appliances. There are two ways for feature separation and extraction.

The first way for feature separation is when the current, power and other parameters are changed because of a switching event. The sampled data before and after the switching event are represented as two different sequences. Equation (2.5) and (2.6) shows two sampled discrete sequences before and after the switching event, represented by  $[a_1]$  and  $[a_2]$  respectively, where  $X_1^1$  to  $X_r^1$  and  $X_1^2$  to  $X_r^2$  are the corresponding values of each sampled data,

$$[a_1] = [X_1^1, X_2^1, X_3^1, \dots, X_r^1] \quad (2.5)$$

$$[a_2] = [X_1^2, X_2^2, X_3^2, \dots, X_r^2] \quad (2.6)$$

hence the change value  $[\Delta a_{2-1}]$  between  $[a_1]$  and  $[a_2]$  can be directly separated by the difference method,

$$[\Delta a_{2-1}] = [a_2 - a_1] = [X_1^2 - X_1^1, X_2^2 - X_2^1, X_3^2 - X_3^1, \dots, X_r^2 - X_r^1] \quad (2.7)$$

where  $[\Delta a_{2-1}]$  represents the variation of  $[a_1]$  changing into  $[a_2]$ . In other words,  $[\Delta a_{2-1}]$  contains features of the switched appliances that changed the current, power and other parameters.

This principle of feature separation, based on the difference method, is simple and less time-consuming. The obtained features are represented in a dataset that contains

massive information. However, one disadvantage of this method is that the sampled data sequence from the residential sites is all in periodic waves. If there is no reference point to fix the phase of sampled data sequences before and after a switching event, then the change value between the two sampled sequences becomes divergent and the authenticity of the data is decreased.

The second way for feature extraction is to use fast Fourier transform (FFT), wavelet transform, S transform or other signal analysis methods[54] [55] to extract electric features such as the appliance power factor, current harmonic components, mean current value, peak current value and power harmonic components. These extracted features can be used to easily identify any particular appliance [56]. However, if several appliances are running simultaneously, the obtained appliance features will be seriously distorted [56]. The resultant features obtained are for all combined appliances and not the individual ones.

Upon comparing the difference method with the extracting parameter method, the former yields a more comprehensive and detailed representation of the appliance features, resulting in a more unique signature. This research will thus employ the difference method. A reference point of the phase will be determined for the two sampled data sequences.

#### ***2.3.4 Appliance Type Identification***

Once the features of switched appliances are extracted or separated, the switched appliance type and working state can then be identified. In NILM, there are three key methods to do this: graph signal processing, pattern recognition and deep training.

##### **2.3.4.1 Identification Model Based on Graph Signal Processing**

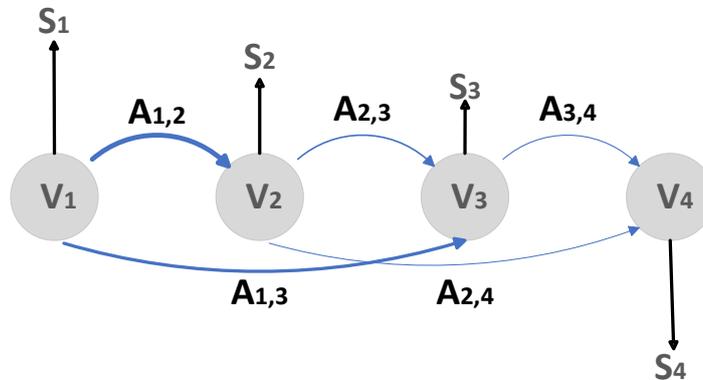
Graph Signal Processing (GSP) [57] is a new data analysis tool for identification and classification. It provides an intuitive means to exhibit the dependency, similarities, or other properties among data elements. It can also represent and process datasets with complex structures. In NILM research, GSP uses a group of nodes that have previously been defined and a weighted adjacency matrix to represent the separated features, and then determine the maximum categories of separated features. It is more

suitable for data classification problems with short training times but establishing appropriate models is challenging.

In GSP, the aggregation of separated features is represented by an unknown dataset  $x$ , which is expressed in the graph of  $G = (V, A)$ , where  $V$  is the set nodes of graph  $G$ , and  $A$  is an  $N \times N$  matrix which is the weighted adjacent matrix of graph  $G$ . Each element  $x_i$  in dataset  $x$  corresponds to one node  $v_i$  in  $V$ . The relationship between node  $v_i$  and  $v_j$  is determined by the element  $x_i$  and  $x_j$ , which is written as:

$$A_{i,j} = \exp \left[ -\frac{(x_i - x_j)^2}{\rho^2} \right] \quad (2.8)$$

where  $\rho$  is a scaling factor,  $x_i$  and  $x_j$  are any two elements in dataset. Based on the relationship between any nodes  $v_i$  and  $v_j$ , a graph is constructed. Fig. 2.4 shows an example of a four-node graph constructed from  $x = [x_1 \ x_2 \ x_3 \ x_4]$ .



**Fig.2. 4 A GSP example with four nodes**

In Fig.2.4, the thickness of the edges in the graph depicts the correlation between nodes.  $S_1, S_2, S_3$  and  $S_4$  denote a set of classification labels. Each edge is similar to a label, and if the difference in features between dataset  $x$  and  $S$  is small,  $S$  is considered piecewise smooth relative to  $G$ . The smallest  $S$  output is indicated by the smoothest graph signal and hence the best classification label. GSP thus provides a powerful, scalable, and flexible data mining and signal processing approach, particularly suitable for data classification when training periods are too short to build any appropriate class models [58]. However, the biggest limitation of using GSP in NILM is that the appliance electrical parameters from the sampled data, such as the working current,

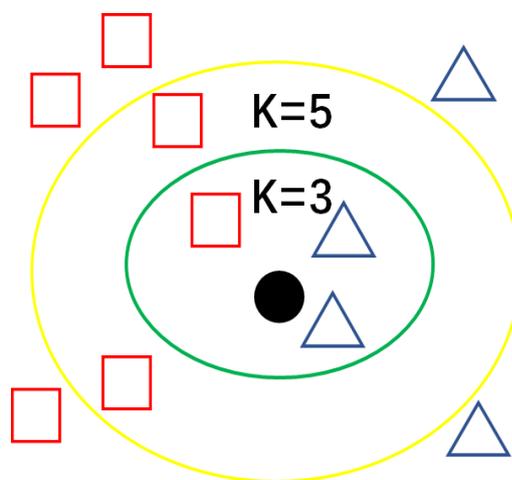
power factor, harmonic components active and reactive power, are transformed into a new matrix during the construction of the graph. In constructing this graph, the physical meaning of the relevant parameters changes, resulting in partial loss of the feature information. Although GSP can identify the type of electrical appliances, it cannot further calculate and quantify the power consumption of the identified appliances.

#### 2.3.4.2 Identification Model Based on Pattern Recognition

Pattern recognition identifies and classifies unknown samples into specific categories based on the feature difference between known and unknown samples. In NILM research, this pattern recognition model must first learn the feature patterns of home appliances to identify the separated features, and then the pattern recognition can obtain the types of home appliances used. The learning process can be done with a supervised learning algorithm or an unsupervised learning algorithm.

For a supervised learning algorithm, prior training on the identification model is necessary to learn the appliance electric features for the appliance identification to be processed online. K-nearest neighbor (KNN) [59, 60] and support vector machine (SVM) [61] are two commonly used supervised learning tools in NILM.

Fig.2.5 shows a basic identification example using the KNN method to illustrate the principle.



**Fig.2. 5 Basic principle of KNN**

In Fig.2.5, the black circle represents the unknown sample while the triangle and rectangle represent the known training samples. The model training process is to learn the features of the known samples. Once the feature learning is completed, the unknown sample is fed into an identification model. If the training samples match the feature of the unknown sample, they will distribute themselves around the latter. When  $K = 3$ , it means that there are 3 training samples in the green circle. As there are more triangles than rectangles, the black circle is thus identified as the triangle category. Conversely, if  $K = 5$ , which means 5 training samples are found. The black circle is identified to belong to the rectangle category because there are more rectangles than triangles (see the yellow circle area). There are three key factors for identification using this KNN method [59]. They are feature difference measurement, size of  $K$ , and classification of the unknown sample based on the  $K$  found training samples. Each of these factors is discussed separately below.

First, the feature difference between the training sample and the unknown sample is calculated. The training samples set is  $A = \{(N_1, y_1), (N_2, y_2), \dots, (N_i, y_i)\}$ , where  $N_i$  is the number label of the training sample, and  $i$  is the number of training samples;  $y_i$  is the  $i$ th training sample in set  $A$ , where the feature aggregation of  $y_i = (F_1, F_2, \dots, F_L)$ , and  $F_1, F_2, \dots, F_L$  are the corresponding values of the component features. Once the training samples set  $A$  is constructed, the training of the KNN model is completed.

The unknown sample is represented by  $B = \{(M_l, x)\}$ , where  $M_l$  is the number label of the unknown sample.  $k$  is one unknown sample,  $x = (G_1, G_2, \dots, G_L)$ ,  $(G_1, G_2, \dots, G_L)$  is the feature aggregation and  $G_1, G_2, \dots, G_L$  are the corresponding values of the component features.

Based on the feature aggregation of the training sample and the unknown sample, the difference between the two samples is commonly calculated by:

$$D(x, y_i) = \left( \sum_{j=1}^L |G_j - F_i^j|^p \right)^{\frac{1}{p}} \quad (2.9)$$

In equation (2.9),  $D(x, y_i)$  is the feature difference between the unknown sample  $k$  and the training sample  $y_i$ .  $G_j$  is the  $j$ th component in feature aggregation of unknown sample  $x$ .  $F_i^j$  is the  $j$ th component in the feature aggregation of the training

sample  $y_i$ .  $L$  is the number of components contained in the feature aggregation of the training sample and the unknown sample. The value of  $p$  is set artificially, which is usually 1 or 2. When  $p = 2$ , this distance is Euclidean distance.

Secondly, the size of  $K$  is determined. As previously established, all training samples are divided into different categories. Each training sample can only belong to one category. The size of  $K$  affects the number of samples that can be used in the identification process, hence it can significantly impact the accuracy of the algorithm used. If the size of  $K$  is small, such as when  $K = 1$ , the unknown sample is related to only one sample and category, thus increasing the identification error. However, if the value of  $K$  is large, the unknown sample can be associated with samples from numerous categories, resulting in an underfit. Generally, for practical applications,  $K$  is a small odd number.

Thirdly, the unknown sample is classified according to the feature difference and the size of  $K$ .  $K$  training samples which are nearest to unknown sample's feature in set  $A = \{(N_1, y_1), (N_2, y_2), \dots, (N_i, y_i)\}$  are found. This is recorded as set  $C = \{(N_1, y_1), (N_2, y_2), \dots, (N_K, y_K)\}$ , and then the category of the unknown sample  $x$  is classified. To keep the classification error rate small, the empirical risk of identification must decrease, and using majority voting can minimize this risk. Therefore, the unknown sample is simply and directly classified by majority voting of  $N_1, N_2 \dots N_K$ .

KNN recently gained prominence as a load monitoring identification algorithm due to its simple principle. However, the identification accuracy is linearly dependent on the feature space dimensionality of training samples [62], it also determines the complexity of model training. When it is applied to detect online residential appliance monitoring, if the feature space dimensionality of training data is large, it is more difficult for online devices to meet such storage needs, yet, if the feature space dimensionality of training data is small, the model training will be insufficient, and the features learning will be uncompleted. The identification accuracy is lower.

Besides the KNN method, SVM is another model that has proven to be successful in various classification scenarios. The SVM's approach is to identify the separation features by finding a set of hyperplanes in high-dimensional space [63], which requires a large support vector (SV) set to ensure identification accuracy but this can lead to an

increase in memory usage and larger computing effort. High dimensionality also demands similar increase in memory usage and computation effort. Hence, SVM may not be suitable for online real-time monitoring but is more suited for off-line data NILM processing.

An unsupervised learning algorithm does not need the prior model training which can directly mine the similar currents used by the electric appliance features. Clustering methods such as K-means, Density-based spatial clustering of applications with noise (DBSCAN) and others are used to extract appliance features for identification, which uses Hilbert transform and other data processing tools to cluster the load with similar signatures [64]. The identification based on an unsupervised learning algorithm can be treated as a blind signal separation problem, which embeds the clustered appliance feature to filter out unwanted features and effectively improves the identification performance [65, 66]. Because unsupervised algorithms achieve identification only through similarity in appliance features, this means that with more types of appliances running in the homes or with multiple appliances always running at the same time, the accuracy is usually lower than that of supervised algorithms, though this is not commonly used in residential sites load monitoring.

#### 2.3.4.3 Identification Model Based on Deep Learning

In recent years, identification based on deep learning (DL) has been gradually applied in this field to help monitoring systems achieve better performance. ANN [67] and BPNN [68] have been gradually applied in NILM. The DL system can analyse massive amounts of data, recognize patterns, and make predictions or decisions without being explicitly programmed to perform these tasks. The DL systems also operate by ‘learning and improving from experience’ [69]. Its algorithm is capable of evaluating whether a prediction and identification is accurate or not. The real capabilities of such a system can be summarized as a process that predicts future events based on past occurrences.

Yang [70] proposed a two-layer neural network with bidirectional recurrent architecture. It combines the HMM (Hidden Markov model) and deep neural network to disaggregate power data, which is tested with the REDD data set [71]. Although the monitoring accuracy is almost 90% based on REDD data, the monitoring accuracy is

greatly decreased under the other data set, because the prior learning experience from the REDD data is useless for a new monitoring scenario. Kelly developed two kinds of deep learning network architecture for non-intrusive load monitoring tasks. The first one is a long short-term memory (LSTM) network [72]; the second one is a denoising auto-encoder (DAE) network [73]. The two deep learning networks were tested using the UK domestic appliance level electricity (UK-dale) data set [74]. The test results show that: the LSTM network has higher identification accuracy for large power appliances, but it is unable to identify low power appliances. The DAE network has higher identification accuracy for low power appliances, but the monitoring accuracy for large power appliances is lower. These two networks cannot keep high monitoring accuracy for all appliances, because the learning strategy and network structure cannot support identification of all types of appliances.

The main disadvantages of monitoring based on the DL system are that: first and foremost, there is no widely accepted a learning strategy and a network structure can be universally used for any monitoring scenario and all household appliances. Designing the specific network structure and learning strategy according to the specific monitoring scenario is the reason behind the highest-accuracy identification result, but a large amount of data is needed to support the learning process, which is an order of magnitude higher than the training data of pattern recognition algorithms, so it is virtually impossible to design a specific network structure and learning strategy for every electricity consumption situation of each load type and each resident home. Secondly, by adopting a DL system in online monitoring, although the ability of learning is more powerful than the traditional pattern recognition method, in order to support the corresponding learning and error correction capability, the hardware requirements of monitoring devices become very high, and the computation process of the whole NILM will be too complex.

Comparing the graph signal processing, pattern recognition and deep learning. The graph signal processing can only identify the appliance type, but it cannot obtain the further data for the power consumption accounting and running characteristics analysing. The identifying accuracy of pattern recognition is lower since the limited online devices storage ability. The deep learning identifying needs to design a specific network structure and learning strategy for different monitoring scenario. Therefore,

when the identification model is completed by combining the pattern recognition and deep training together, the shortcomings of each method can be mutually compensated. The combining of pattern recognition and deep training can improve the accuracy of pattern recognition identifying and use the fixed learning strategy for majority monitoring scenario.

## **2.4 Proposed Monitoring Approach**

After reviewing the data sampling, event detection, feature separation and feature identification methods, a successful online NILM for residential appliances can be achieved by leveraging the strengths of relevant methods and overcoming their limitations. To complete the monitoring with low computational complexity, separate the integral appliance features, and accurately identify the appliance type with a short response time, the monitoring method is designed as follows: firstly, the current data is sampled with high frequency at the home electricity supply entry point. Secondly, the sampled data is processed through these three steps to obtain the appliance type, power  $P_m$ , and state  $a_m$ . 1) the Heuristic method is used to detect appliance switching events, 2) the difference method separates the appliance features, 3) the KNN method combines with BPNN to identify the appliance type. The reason why choosing corresponding methods and what improvements have been made to the shortcomings of the chosen methods are introduced briefly.

### ***2.4.1 Switching Event Detecting with Heuristic Method***

The Heuristic method is used to detect appliance switching events. Due to its simple calculation process and reduced complexity, this method minimizes the use of online computing power and shortens the response time.

To overcome the issue of low detecting sensitivity for small power appliances, the threshold value is combined with an additional judging condition such as evaluating the change of all sampled data rms value between two observation periods. This enhances detection noise immunity and improves detection sensitivity. For this thesis, the auxiliary time meter is eliminated to improve the feasibility of Heuristic detection in NILM. The specific detection process is discussed in detail in Section 4.3.

### ***2.4.2 Separating Current Feature with Difference Method***

The difference method is chosen as the method used to separate the appliance's current features for two reasons. Firstly, the difference method can maintain the authenticity of data and ensure the separated data contains enough information for appliance type identification. Secondly, the system noise will always exist in the practical power system, and the difference method has a stronger anti-noise ability to separate electrical features, without noise reduction on the original data.

As for the issue of having no reference point to fix the phase of sampled data sequences when using this method, it can be resolved by using the voltage and current signals together. The voltage datasets as the reference to sample and separate current data. This process is further elaborated in Section 4.4.

### ***2.4.3 Accurate Identification of Appliances with Similar Feature***

The identification is achieved by combining the KNN method with the DL networks to optimize both approaches' identification accuracy and response time requirements. The identification is fundamentally based on the KNN method which shortens the overall monitoring response time. The identification result is then modified and corrected using the appliance operation pattern which is regressed by BPNN to improve its monitoring accuracy.

A KNN identification model is reconstructed, and the feature space dimensionality of training data is selected and compressed based on the prior appliance features analysis to minimize the effect of insufficient online storage capacity. The specific KNN model building process is set out in Section 5.2. To overcome the issue of poor universality learning strategy and high calculation complexity in BPNN application, this thesis uses BPNN to regress the appliance operation pattern, so that the learning strategy for obtaining appliance operation pattern can be applied to the majority monitoring scenario. Furthermore, the BPNN is only used as an offline technology to process the appliance operation feature without increasing the complexity of the monitoring method. This process of BPNN training and operation pattern regressing is detailed in Section 5.3.

## 2.5 Summary

Using NILM for demand-side power management received noticeable interest due to its significant advantages over intrusive monitoring. This chapter reviewed each step towards achieving NILM. This includes data acquisition, appliance switching event detection, appliance feature separation, and appliance identification. The strengths and limitations of different solutions in each step were also analysed. Although low-frequency data acquisition is the widely used technology for NILM, high-frequency data acquisition has gained significant attention due to its attractive advantages of preserving the appliance features. Compared to the slow processing of mutation point detection, Heuristic detection is found to be more suitable for NILM. Using a pattern recognition method to identify the appliance type, the appliance's electric features and operation features are combined to improve the monitoring accuracy. This review showed that as each step of NILM is highly related to appliance features, a comprehensive understanding of this parameter is fundamental for successful appliance switching event detection, appliance feature separation, and appliance identification model build. In the next chapter, a comprehensive comparison and combination of available appliance features will thus be first explored.

## **Chapter 3 Appliance Features Analysis and Extraction for NILM**

### **3.1 Introduction**

Extensive research in NILM focuses on identifying an effective set of features with unique appliance signatures that can discriminate different loads. While many appliance features have been reported in research, only a limited subset of them were used. There is also an absence of research that carried out a systematic comparison and combination of the features. Systematic selection of various electric features to discriminate appliances was also not present. This chapter shall analyse some appliance features and compare them systematically.

Firstly, the different electric features are carefully described and reviewed, then the different electric features are extracted individually from public data dataset. The features with obvious peculiarity are then selected and combined as the key ones for online NILM processing in Section 3.2. Next, the electric data of eight typical appliances are sampled in a real residential power consumption environment to extract the selected appliance features. The results are shown in section 3.3. Lastly, the non-electric features of appliances to overcome the overlapping electric features are discussed in Section 3.4.

### **3.2 Researching of Household Appliance Electric Feature**

The features that can be extracted from the measurement data depend on the sampling rate [75]. All features can be categorized into steady and transient states, depending on the state of the measured waveform they represent [76]. Identifying and distinguishing the different appliance features is the most important step in the monitoring process. However, not all features can be recognized and distinguished in NILM. In contrast to other direct monitoring methods, the objective of this section is to choose a subset of features that can either better or provide comparable discriminative performance to using all of them. A comprehensive list of steady and transient state features, including references to their extraction steps, is outlined in Table 3.1.

**Table 3. 1 Notation and description of extracted features**

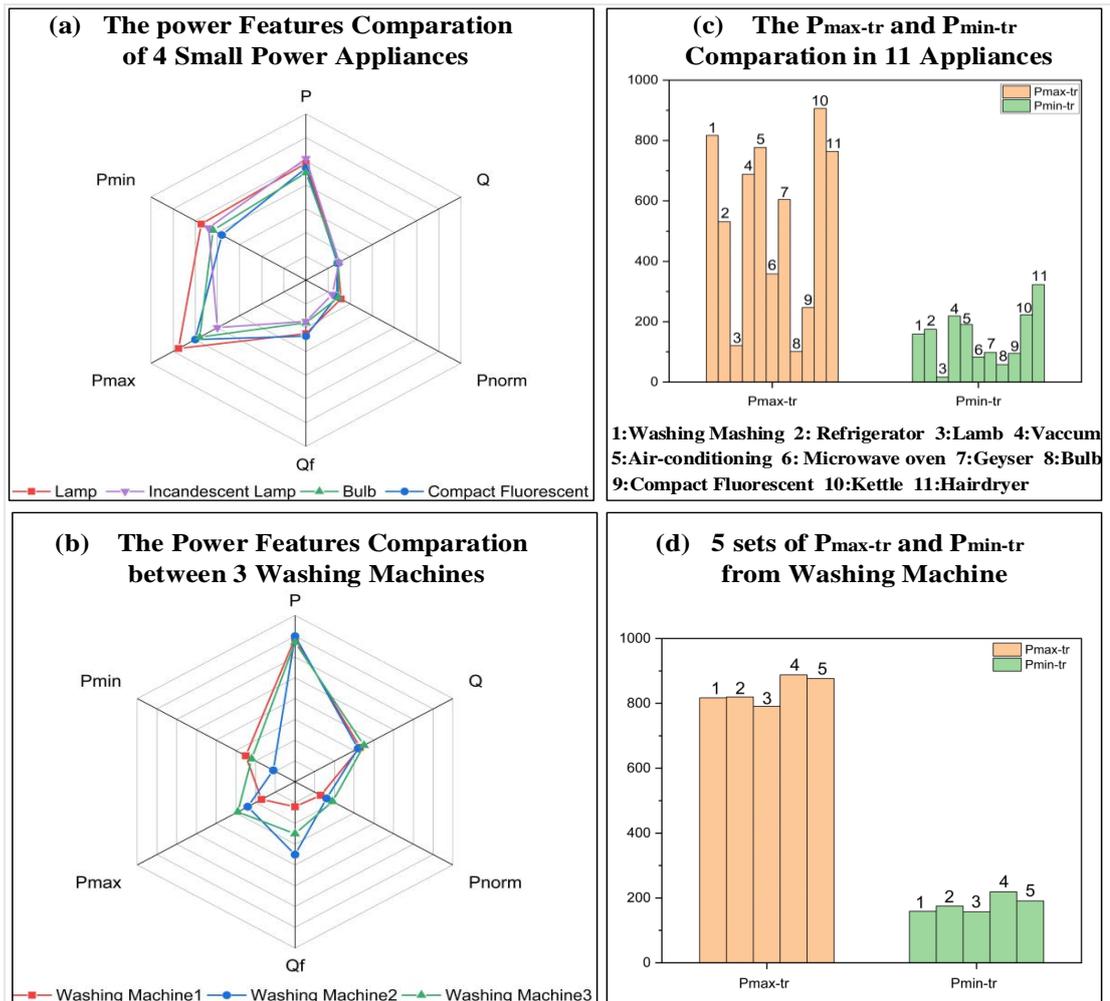
Steady State Features	$P$	Real power[77]
	$P_{norm}$	Normalized real power[78]
	$Q$	Reactive power[79]
	$Q_f$	Reactive power based on Fryze's formula
	$S$	Apparent power[80]
	$I_{rms}$	Current root mean square
	$I_{f-rms}$	Nonactive current root mean square
	$I_{har(j)}$	$J$ th current harmonic coefficient
	$I_{f-har(j)}$	$J$ th nonactive current harmonic coefficient
	$V_{har(j)}$	$J$ th voltage harmonic coefficient
	$I_{THD}$	Total harmonic distortion of current[81]
	$I_{fTHD}$	Total harmonic distortion of nonactive current
	$V_{THD}$	Total harmonic distortion of voltage
	$Asymmetry$	Asymmetry measure of asymmetry in V-I trajectory[82]
$Intersections$	Number of intersections in V-I trajectory[82]	
$Area$	Enclosed area by VI trajectory with consideration of trajectory direction[82]	
$Net area$	Net area enclosed by VI trajectory without consideration of its direction[82]	
$Curvature$	Measure of distortion of mean line of VI trajectory from a straight line[82]	
$Slop$	Slope of the middle segment of V-I trajectory[82]	
Transient State Features	$Wd(i)$	Energy of detail wavelet coefficients at $i$ th scale[83] [84]
	$W_{max-idx}$	Index of the maximum energy wavelet coefficient[83] [84]
	$W_{max}$	Maximum value of the wavelet coefficient
	$I_{max-tr}$	Maximum value of the transient current[85]
	$I_{max-tr-idx}$	Location of maximum transient current[85]
	$I_{min-tr}$	Minimum value of the transient current
	$I_{min-tr-idx}$	Location of minimum transient current
	$diff_{I-idx}$	Difference between maximum and minimum values of transient current
	$P_{max-tr}$	Maximum value of the transient power
	$P_{max-tr-idx}$	Location of maximum transient power[86]
	$P_{min-tr}$	Minimum value of the transient power
	$P_{max-tr-idx}$	Location of minimum transient power[69]
$I_{peak-num-tr}$	no. of local maximums of transient current[87]	

In Table 3.1, the 32 appliance features are commonly employed for load identification in NILM research. They can be classified into the four parameters of power, current, harmonic and wavelet:

Power category: Real power ( $P$ ) and reactive power ( $Q$ ) are the most commonly used steady state features to identify appliances that have only two running states (ON/OFF) and high-power consumption. However, it is difficult to discriminate the low-power consumption appliances using only these  $P$  or  $Q$  features due to the significant overlap in the corresponding feature space. To overcome these overlaps during the identification process, the  $P_{norm}$  (real power after normalizing), the  $Q_f$  (reactive power after Fryze's formula), the  $P_{max}$  (maximum steady power) and  $P_{min}$  (minimum steady power) are used instead and they are extracted as a feature set. However, these four features are only effective in identifying appliances with significant spikes in their power draw which is followed by slower changing variations. Comparing these six steady state power features ( $P$ ,  $Q$ ,  $P_{norm}$ ,  $Q_f$ ,  $P_{max}$ , and  $P_{min}$ ) with instantaneous power features such as the  $P_{max-tr}$  (maximum value of transient power) and  $P_{min-tr}$  (minimum value of transient power), the instantaneous power features are higher in their uniqueness, as transient features can overcome overlapping issues to identify most appliances.

According to the above-mentioned analyses, the eight power features are extracted individually from the public dataset. Fig.3.1 displays the extracting result, it compares the feature differences between appliances of the same type and the feature similarities between appliances of different types. In Fig.3.1(a), the low-power consumption appliances such as Lamp, Bulbs, Compact Fluorescent and Incandescent Lamp have significant overlaps in  $P$  and  $Q$  features. In Fig.3.1(b), although all the  $P_{norm}$ ,  $Q_f$ ,  $P_{max}$ , and  $P_{min}$  features are extracted from the Washing Machine electric data, there is a significant difference in the corresponding values of the same parameter. Therefore, these six power features of  $P$ ,  $Q$ ,  $P_{norm}$ ,  $Q_f$ ,  $P_{max}$ , and  $P_{min}$  may not be suitable for further appliance identification. Conversely, Fig.3.1(c) shows that the instantaneous power feature can distinguish the 11 appliances from different categories. In Fig.3.1(d), the similarity of the instantaneous power feature is also high in the Washing Machine category. Therefore, instantaneous power features of appliances have a better distinguishable nature in the power category, and it will be extracted for

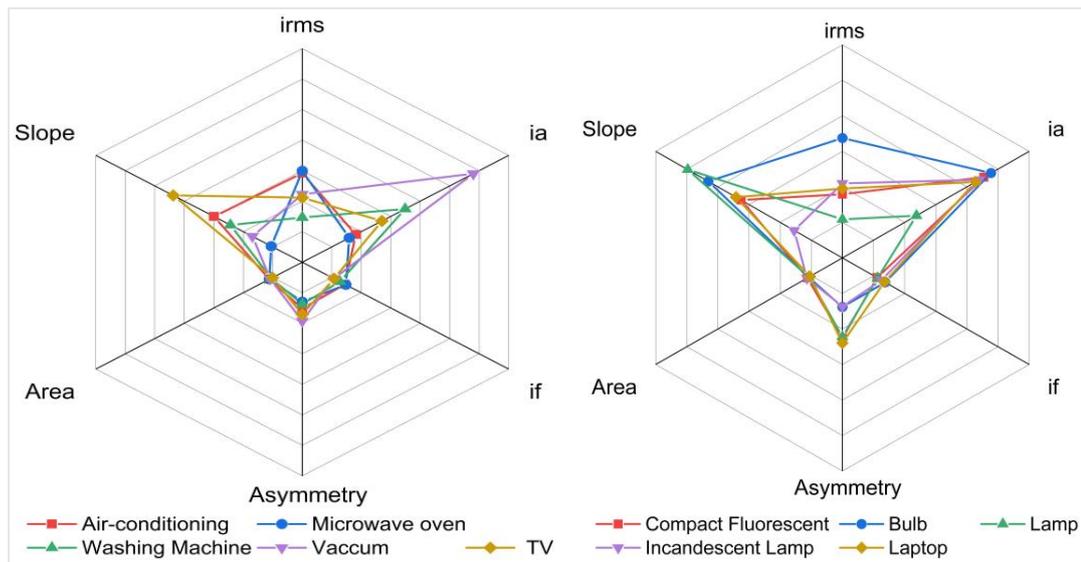
further NILM processing.



**Fig.3. 1 Power features comparing between different appliances**

Current category: With more non-linear appliances being used, the steady and transient working currents of these appliances get more non-linear and non-sinusoidal, so the current features become more multi-dimensional. Firstly, the differences in  $i(t)$  (the steady working current) and  $i_{rms}$  (the root mean square of current) of different appliances become distinct enough to directly differentiate the appliances. Secondly, the steady working current  $i(t)$  can be decomposed into active current  $i_a(t)$  and non-active current  $i_f(t)$ . The similarity between  $i_a(t)$  draws in different appliances (with similar power levels) is lower compared to their power waveforms. Thirdly, plotting the steady current signal against the steady voltage signal for an appliance V-I curve makes the shape-based feature extraction computationally efficient, making this

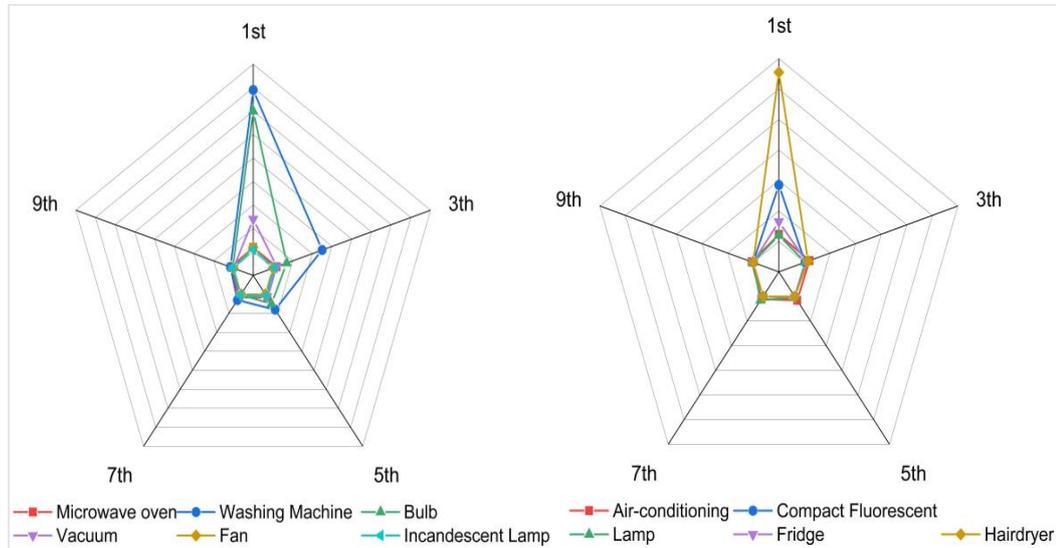
current feature approach more robust. However, the voltage signal is stable in residential sides, the six parameters of V-I trajectory (Asymmetry, Intersections, Area, Net Area, Curvature and Slope) are dependent on the steady current signal. Based on the twelve current features analyses, the appliances'  $i_{rms}$ , active current  $i_a(t)$ , nonactive current  $i_f(t)$  and V-I trajectory parameters are extracted individually from the public dataset. The results are displayed in Fig.3.2, which shows a reduced overlap between each feature, making the corresponding current features more identifiable. Therefore, the appliance steady working current is vital for appliance identification and further feature extraction. A steady working current will thus be used as the primary feature for further NILM processing.



**Fig.3. 2 Appliances current features comparing**

Harmonic category: Harmonic contents of current waveforms (obtained by Fast Fourier Transform of high frequency measurements) not only can discriminate non-linear and multi-state appliances, but also make current features more recognizable. The magnitude and phase angle of the  $I_{har-j}$  ( $J$ th current harmonic) of steady state current are unique signatures for appliance identification. Furthermore, the 1<sup>st</sup> to 4<sup>th</sup> harmonic can specifically improve the discrimination between appliances, making them a key coefficient for non-linear appliance identification. The harmonic components along with total harmonic distortion ( $I-THD$ ) of current waveforms pave the way for the identification of non-linear and variable-working state appliances. The

appliances' steady current harmonic components are extracted from the public dataset, and the result is shown in Fig.3.3, it demonstrates a clear and distinguishable pattern. The overlaps between different appliance harmonic component sets are lesser. Therefore, the appliance harmonic contents of steady working current can be extracted for further application.



**Fig.3. 3 Appliances harmonic features comparing**

Wavelet category: A wavelet transform can decompose a signal into time and scale using wavelets that have adaptable scale properties. Continuous wavelet transforms (CWT) of a signal can construct a new load signature that is totally different from the original signal. It decomposes switching voltage transients and is thus recommended in reference[88] as a promising approach to extracting transient features in NILM. To avoid computational complexity using CWT, a Discrete Wavelet Transform (DWT) is used as an alternative to extract features for NILM applications. Reference[89] applied DWT of the turn-on current transient as a feature set and showed advantages over CWT in the transient analysis of loads. Wavelet features achieve high discrimination accuracy due to the higher difference between various appliances in this wavelet category. But the wavelet features among appliances of the same category are also very different, making them unsuitable for real-time applications. Hence, wavelet features will not be extracted and used for further identification.

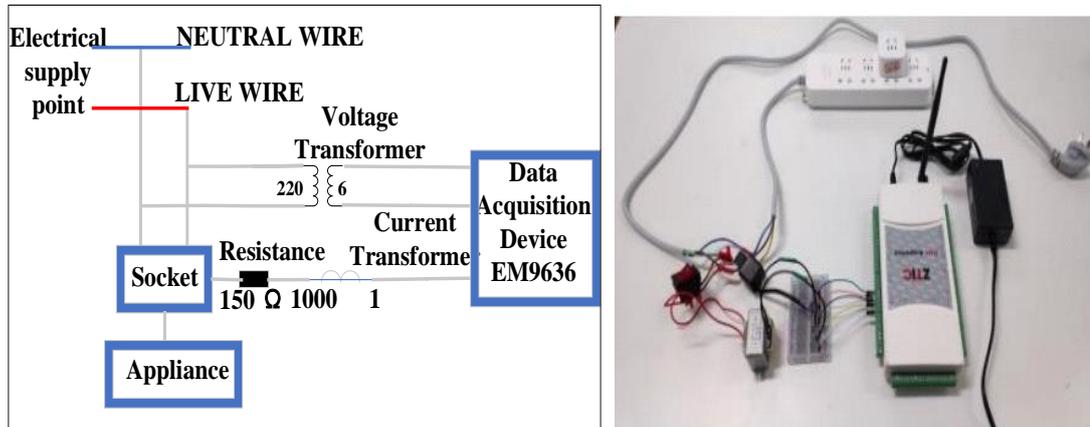
While power, current, harmonic and wavelet features are effective in appliance type identification when used individually, the combination of features in different categories can further improve the uniqueness of these features. Millers [90] and Niesche [91] used the regression coefficient of the appliance working current to complement the  $P$ - $Q$  features. Lu[52] used the total harmonic distortion of current waveforms along with  $P$  and  $Q$  for load identification. Ahmadi[92] used FFT of voltage noise to define appliance signatures and later extended their work to consider Fourier features of the electromagnetic interference signals in the 36-500 kHz range [93]. Kong [94] used the frequency and amplitude of the dominant peaks in the smoothed cepstrum of the voltage and power signal as appliance features to distinguish ON/OFF appliances. The cepstrum is defined as the inverse Fourier transform of the logarithm of the spectrum of a signal [94]. Sultanem [95] combined the current harmonics with low frequency-based power features and voltage wavelets. However, this combination is only used to distinguish appliances in the electric dimension, and for some appliances, the electric features will always be similar. To improve the degree of appliance differentiation, it is vital to take non-electric features into account. The appliance non-electric features such as the time and frequency of the appliance usage are used to complement the feature combination set when the corresponding appliances cannot be identified. Therefore, the continuous power signal with a set of discrete pulses which are about human behaviour information (time of appliance switching and duration of appliance operation) will be also extracted for the identification.

Effective and universal appliance electric features will be extracted in a real residential power consumption environment. The extraction results will be discussed in the next Section 3.3, which include active power, reactive power, working current, working current harmonic components, instantaneous power and non-electric feature related to appliance operation.

### **3.3 Appliance Electric Feature Extraction**

In this section, voltage and current data of different appliances were sampled using the independent data acquisition device. The various electric features were then extracted from this acquired data. The data acquisition device is shown in Fig.3.4. The

parameters of the data acquisition device are as follows: the main device used is the EM9636B high-precision data acquisition card, which can realize the analogue-to-digital data conversion and support 16-channel differential parallel sampling; the ratio of the voltage transformer in the data acquisition device is 220:6; the ratio of the current transformer is 1000:1; the resistance of the current transformer is 150  $\Omega$ ; the sampling frequency of the data acquisition device is set to 10kHz; the main frequency of the power system is set at 50Hz; the voltage of the power system is set at 220V.

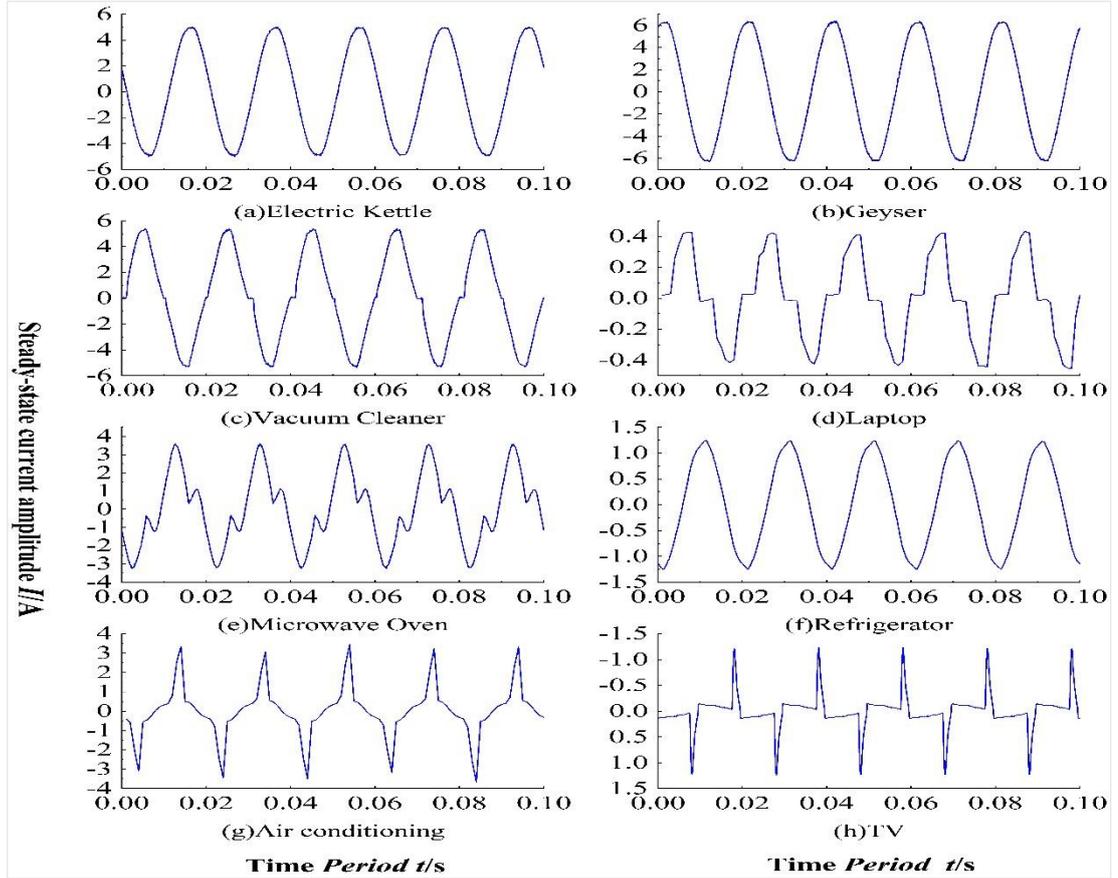


**Fig.3. 4 Schematic diagram of data acquisition device**

The current and voltage data of the air conditioning, refrigerator, geyser, electric kettle, microwave oven, TV, vacuum cleaner and laptop were sampled. During this data acquisition process, the socket as shown in Fig.3.4 is connected to the main power supply point of the house, and the appliance is connected to the socket to enable the collection of the current and voltage data of the independent operation appliance.

### ***3.3.1 Appliance Steady State Working Current***

Using the current transformer developed in the data acquisition device, the appliance steady state working current data can be extracted directly. The sampled current data sequence of appliances is plotted to show the appliance working current waveform, in which an entire sequence includes 1000 sampled data within 0.1 seconds. Fig.3.5 shows plots of sample data of the eight appliances' working current waveforms and their steady state working currents.



**Fig.3. 5 the steady state working currents of appliances**

The steady state working current for different appliances is distinctive. The steady state working current of the electric kettle and geyser is obviously in the form of a sine wave, while the working currents of the microwave oven, vacuum cleaner, laptop, air conditioning and refrigerator were distorted differently and formed different shapes. To better describe the distinct steady-state working current of each appliance, the root mean square ( $I_{rms}$ ) and amplitude ( $I_p$ ) of the steady-state working current are calculated as:

$$I_{rms} = \sqrt{\frac{1}{N} \sum_{k=0}^N i(k)^2} \quad (3.1)$$

$$I_p = \max(i(k)), 0 \leq k \leq N \quad (3.2)$$

where  $i(k)$  is  $k$ th sampled current data,  $N$  is the number of the sampled data in one current sequence, amplitude  $I_p$  is the peak current in one current sequence, the crest

coefficient ( $I_{CF}$ ) is the ratio of the peak current and  $I_{rms}$  value. Using the equations (3.1) and (3.2), the result is as shown in Table 3.2:

**Table 3. 2 Feature values of steady state current for each appliance**

Appliance	rms/(A)	Amplitude/(A)	Crest Coefficient
Electric Kettle	3.5376	5.0000	1.4134
Geysers	4.5021	6.4000	1.4216
Air Conditioning	8.6570	10.4288	1.2047
Refrigerator	4.8666	5.2500	1.4424
Microwave oven	1.9510	3.6000	1.8452
Vacuum cleaner	3.4555	5.4000	1.5627
Laptop	0.2692	0.4577	1.7002
TV	0.6782	0.5273	1.9381

According to Table 3.2, it is clear that the steady state working currents of specific domestic appliances have certain regularities, such as their crest coefficients, that can be used to realize appliance identification. The crest coefficient of resistive appliances is closer to 1.4. Furthermore, the higher the non-resistive component of the appliance, the more the crest coefficient deviates from 1.4.

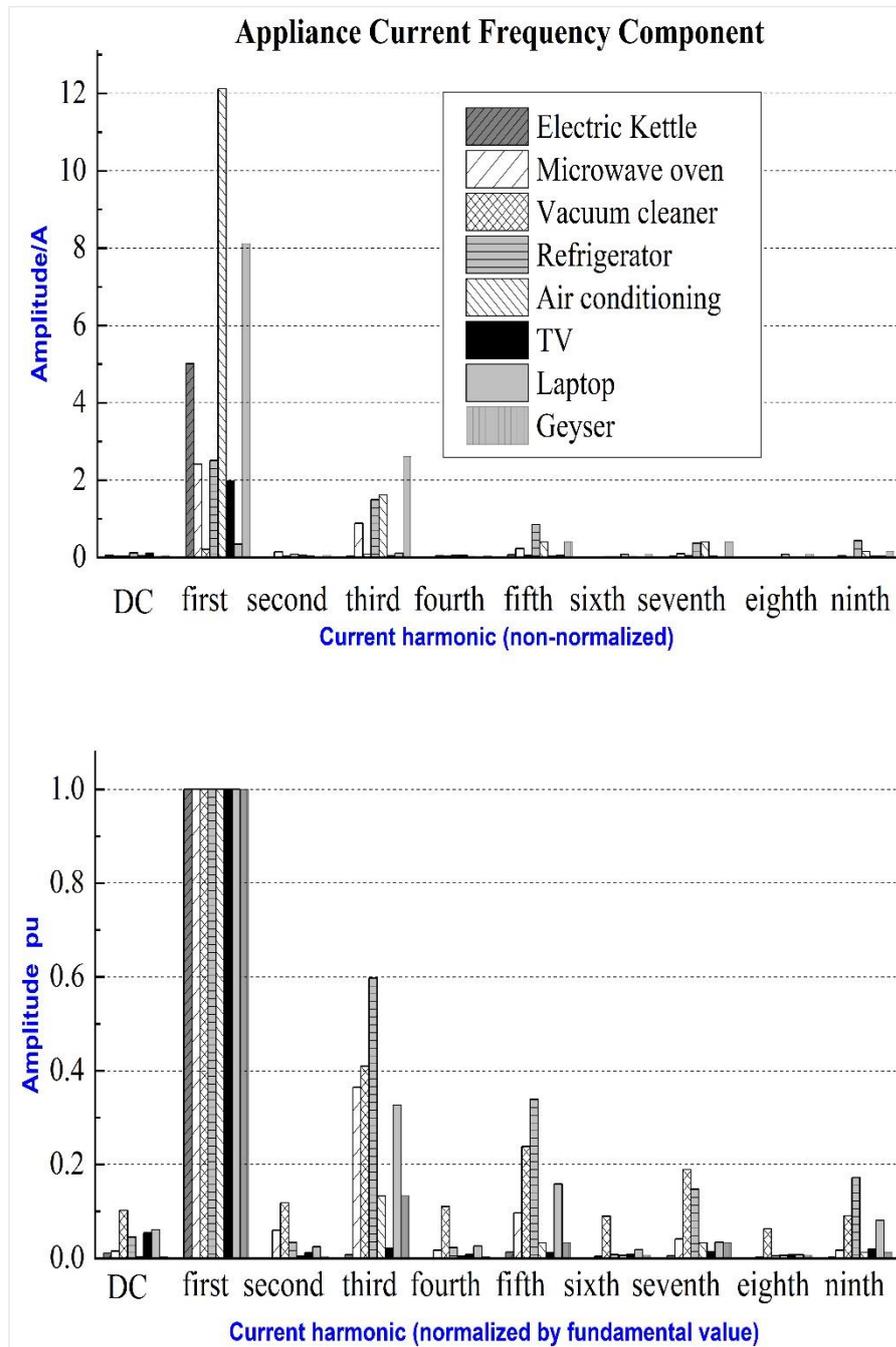
### 3.3.2 Appliance Working Current Harmonics

Transferring the current data into the frequency domain can make the current features even more distinct. The sampled current data sequence can be processed to obtain its harmonics in frequency domain using Fourier series as shown below:

$$I(t) = a_{k0} + a_{k1} \sin(\omega t + \theta_{k1}) + a_{k2} \sin(2\omega t + \theta_{k2}) + \dots + a_{ki} \sin(i\omega t + \theta_{ki}) \quad (3.3)$$

where,  $I(t)$  is the obtained current signal,  $a_{k0}$  is the DC component,  $a_{k1}$  represents the fundamental component of current signal,  $\theta_{k1}$  is phase angle of the fundamental signal,  $a_{ki}$  represents the  $i$ th harmonic component of current signal,  $\theta_{ki}$  is phase angle of the  $i$ th harmonic signal. Once the frequency spectrum is obtained, the fundamental

component (first harmonic) is taken as the baseline to normalize the other harmonic components. The appliance working current harmonics are shown in Fig.3.6.



**Fig.3. 6 Current harmonics of domestic appliance load**

In addition, the total harmonic distortion (THD) value of each appliance can also be calculated based on the current harmonic components. The calculation equation is represented as follows:

$$THD = \sqrt{\sum_{k=2}^N \left(\frac{i(k)}{i(1)}\right)^2} \quad (3.4)$$

where  $i(1)$  is the fundamental component of the appliance working current,  $i(k)$  is the  $k$ th harmonic of working current. According to equation (3.5), the THD value of each appliance load is calculated, and their results are presented in Table 3.3.

**Table 3.3 THD value for each appliance**

<b>Appliance</b>	<b>THD</b>
Microwave oven	38%
Vacuum cleaner	47%
Air conditioning	14%
Laptop	37%
TV	14%
Electric kettle	2%
Geyser	7%
Refrigerator	13%

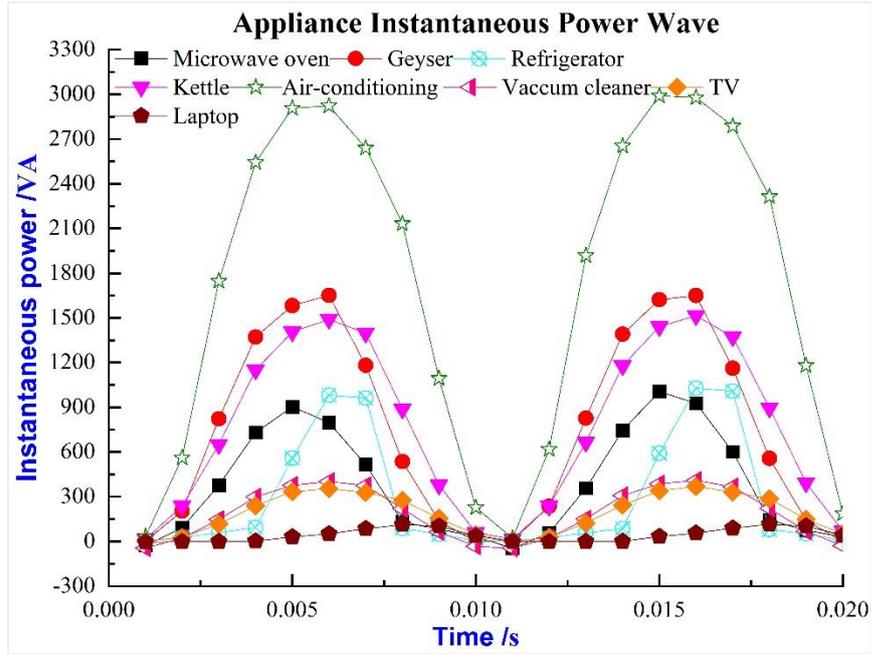
Comparing the working current harmonics of different electrical appliances, it is evident that the distortion rate of the vacuum cleaner is the highest, with a THD value of up to 47%. The THD values of microwave oven and laptop are both approximately 38%. The difference in amplitude between each harmonic is significant, making the distinction between them clear.

### ***3.3.3 Appliance Transient and Steady Power Features***

The appliance instantaneous power can be derived from the appliance working current and voltage data, using the following equation:

$$p(k) = v(k)i(k) \quad (3.5)$$

where  $p(k)$  represents the instantaneous power,  $v(k)$  is the  $k$ th sampled voltage data,  $i(k)$  is  $k$ th sampled current data,  $1 \leq k \leq N$ ,  $N$  is the number of the sampled data in one period of 0.02 second. Fig.3.7 shows the instantaneous power curve of a typical load. The fundamental characteristic of the instantaneous power curves of different appliances is the periodicity, but the amplitude of each waveform is different. The overlapping between these instantaneous power curves is unserious, which can be treated as a unique feature.



**Fig.3. 7 Instantaneous power curve of typical load**

To provide a clearer description of the instantaneous power of each appliance, the power amplitude ( $P_p$ ) and the doublet area ( $S_{pw}$ ) are introduced to re-represent instantaneous power feature, which are defined as follows:

$$P_p = \max(|p(k)|), 1 \leq k \leq N \quad (3.6)$$

$$S_{pw} = \sum_{k=1}^N |p(k)| \quad (3.7)$$

The amplitude and doublet area of appliance instantaneous power are calculated as the main parameters, as listed in Table 3.4:

**Table 3. 4 Instantaneous power feature of each appliance**

Appliance	Doublet area/(J)	Amplitude/(VA)
Microwave oven	75.42	1032.46
Vacuum cleaner	116.41	1632.95
Air conditioning	98.68	1067.56
Laptop	5.35	79.59
TV	36.61	409.43
Electric kettle	31.03	363.52
Geysers	154.63	1524.63
Refrigerator	135.46	1419.57

It can be seen from Table 3.4 that the instantaneous power parameters of different appliances are significantly different. The maximum value of the vacuum cleaner is 1632.95VA, but the amplitude of the laptop is only 79.59VA. These large differences can thus be extracted to realize load identification.

Compared with the appliance instantaneous power, appliance steady state power is not suitable for appliance type identification because of the serious overlapping, but steady state power is a vital parameter to calculate power consumption and analyse the running state characteristic, such as the active ( $P$ ) and reactive ( $Q$ ) power which are calculated as follows:

$$P = \sum_{k=0}^{\infty} P_k = \sum_{k=0}^{\infty} V_k I_k \cos(\varphi_k) \quad (3.8)$$

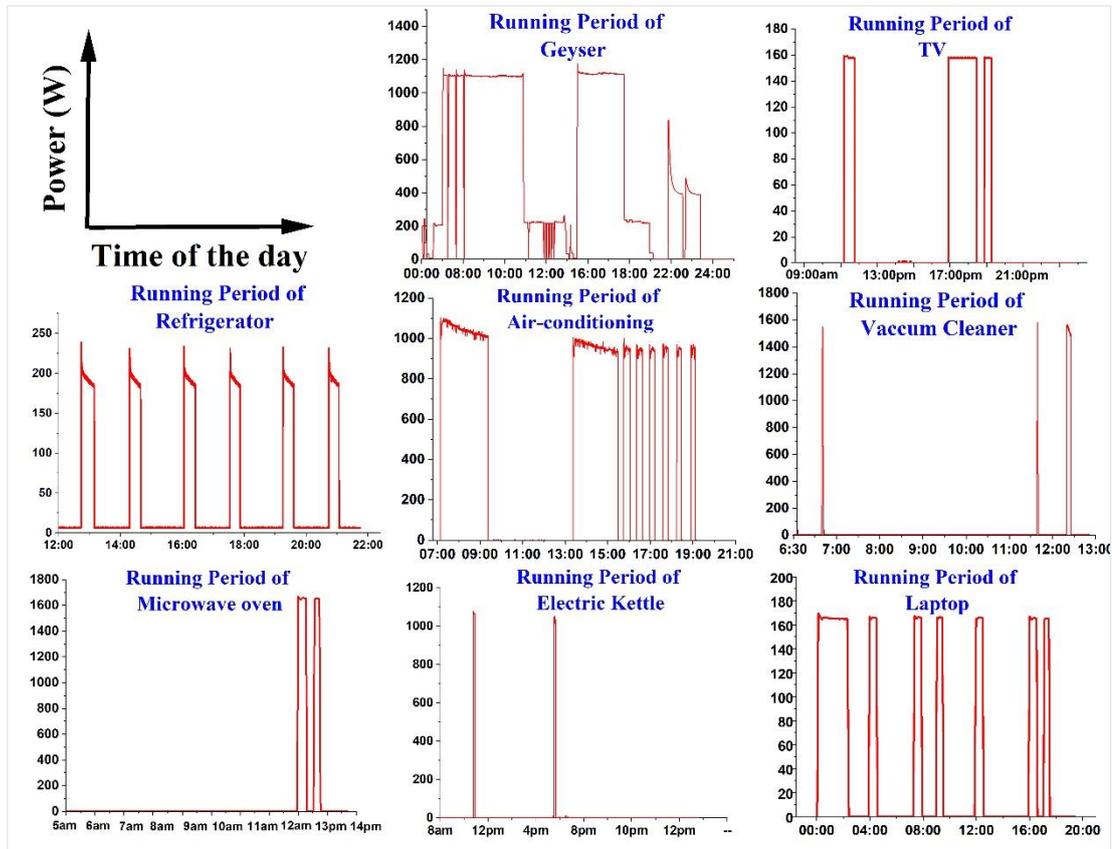
$$Q = \sum_{k=0}^{\infty} Q_k = \sum_{k=0}^{\infty} V_k I_k \sin(\varphi_k) \quad (3.9)$$

where  $V_k$  and  $I_k$  are the  $k$ th harmonic value of voltage and the  $k$ th current respectively of a normal working appliance.  $\varphi_k$  is the power factor angle, and  $k$  represents the harmonic number.

### 3.4 Non-electric Feature Introduction

Despite all the above efforts to select and extract appliance electric features, it remains challenging to generate a unique appliance signature capable of discriminating between various loads. As for the non-electric features, these features are dependent

on the customer's usage habits, which is another almost unique appliance feature, such as the appliance running regularity and operation patterns. Fig.3.8 shows the running duration of different appliances for one customer in a day, the running of eight appliances is highly distinct, which illustrates that the operation pattern of each appliance is a unique and valid feature.



**Fig.3. 8 Appliances running period**

In Fig.3.8, each appliance exhibits a unique operation period curve. Compared to other appliances, the running periods of the refrigerator are relatively fixed, and the duration of the running time are relatively constant. In other words, the operation pattern for the refrigerator is uniform throughout the day. Therefore, even if the electric feature of the refrigerator is similar to other appliances, its operation pattern remains unique. In contrast, the microwave oven and other kitchen electrical appliances are usually switched on/off and running between 11 pm to 1 pm. Though their running periods are relatively stable, their operation patterns will require a longer observation period to obtain. The running time of the air-conditioning, TV, laptop and geyser is not

constant throughout the day, so the operation patterns vary and will also require a longer observation period to obtain operational regularity. Based on the different usage characteristics of the eight appliances, the appliances were thus classified into three categories. Each appliance was observed over different periods to obtain its corresponding operation pattern. Once obtains the appliance operation pattern, it will be unique features, which can help to improve load identification accuracy, especially for loads with similar electrical characteristics [56].

1. Self-running appliance: This type of appliance keeps running intermittently throughout the day such refrigerator and geyser. The appliance running pattern is fixed, and the switching probability is unchanged. The number of switchings and the corresponding switching time recorded within one day can reflect a non-electric feature.
2. Manually switch on/off appliance: These appliances running is determined by usage habits, such as TV, oven, and laptop, which is highly related to personal work and rest. The switching probability is related to usage habit, so the number of switchings and the corresponding running time need to record over a long period. The longer the recorded period, the more precise its non-electric feature is extracted.
3. Seasonal appliance: These appliances switching probability varies in different seasons, such as air-conditioning and heating appliances when these non-electric features need to be extracted individually for a different season.

### **3.5 Summary**

This chapter concludes the review of appliance features for the purpose of load identification. Through systematic analysis and comparison of a wide variety of features, some of the more distinguishable and relevant features were selected for use in NILM. These features were then extracted from actual residential appliances.

Once the features of different appliance are effectively obtained and extracted, the prior research of NILM is completed. These features will then be used to train and construct the monitoring model, for the monitoring of individual appliances in a non-

intrusive way. This monitoring process includes appliance switching event detection, appliance feature separation and appliance type identification, and they will be discussed in the next two chapters.

## **Chapter 4 Appliance Switching Event Detection and Working Current separation**

### **4.1 Introduction**

In the introduction in Chapter 1 and the review in Chapter 2, it was explained how an individual appliance can be monitored through the analysis of sampled data at the electricity supply point. This process includes appliance switching event detection, appliance electric feature separation, and the appliance type identification. This chapter aims to propose an approach for the switching event detection and the appliance working current separation.

Firstly, the characteristics of the household appliance running and switching are analysed. The relationship between the appliance working current and the house current will be further discussed in Section 4.2.

Secondly, the switching event of appliance is detected using the improved Heuristic detection method, which improves the efficiency of the original approach by reducing the analysis of duplicate data. Furthermore, the improved Heuristic method eliminates the use of an auxiliary time meter and adapts the high noise environments. This process will be elaborated in Section 4.3.

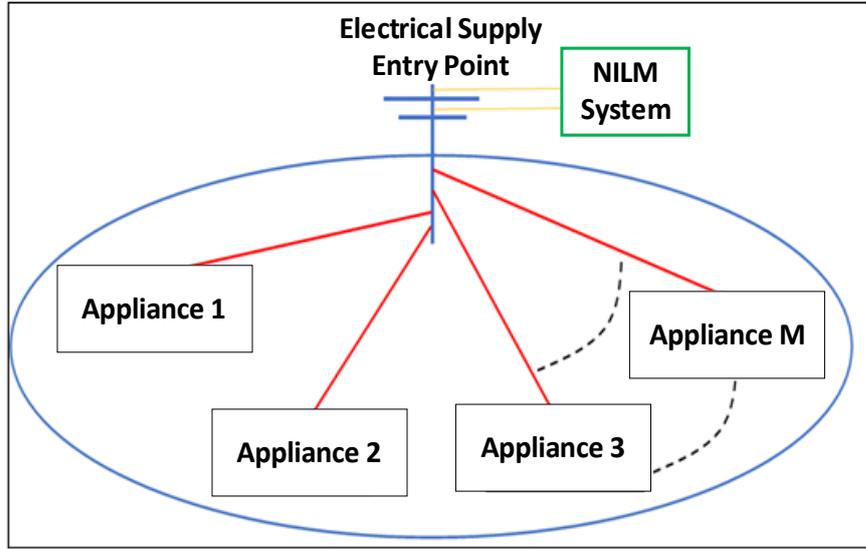
Thirdly, the current feature of the switched appliance is separated after the switching event using the difference method. The limitation of the difference method is resolved, and the phase angle of the separated current can be obtained. More details are provided in Section 4.4.

Lastly, the proposed method for switching event detection and appliance current separation is tested using the UK-DALE [57] dataset. The results are listed and discussed in Section 4.5.

### **4.2 The Character of Running Household Appliances**

Fig.4.1 illustrates the relationship between the various appliances within a house. All the appliances are connected in parallel, with the running of each appliance independent to the others. Once the appliance is switched on and starts running. The

appliance power is supplied by the electricity supply point. The power at the electricity supply point is the sum of all the running appliance powers.



**Fig.4. 1 Appliance's connection diagram**

In Fig.4.1, the house current at the electricity supply point is also the sum of all operating appliances' working currents. Therefore, the house current can be expressed by equation (4.1).

$$I(t) = \sum_{k=1}^M I_k(t) + I_n(t) \quad (4.1)$$

where,  $I(t)$  represents the house current which increases or decreases depending on whether the appliance is switched ON or OFF.  $I_k(t)$  represents the working current of an appliance.  $M$  represents the number of running appliances. The  $I_n(t)$  represents the noise causing distortion to the appliance working current and fluctuation to the house current. There are two noise sources. The first is from the appliance abnormal transients and appliance running fluctuations. As this causes minimal change to the appliance working current, power and other electric parameters, this source is considered negligible. The second noise source comes from the large harmonic intrusion from the external environment. This can cause fluctuation to the house current.

If the harmonic intrusion is not significant, the  $\sum_{k=1}^M I_k(t)$  in equation (4.1)

remains stable and unchanged prior to the appliance switching event. Once the appliance is switched on, the house current will change to a new stable state, and this steady state will maintain for a period as most household appliances almost never switch off once they are turned on. Therefore, the house current  $I(t)$  after a switching event can be expressed using the following equation:

$$I(t) = \sum_{k=1}^M I_k(t) + I_{M+1}(t) + I_n(t) = I_b(t) + I_a(t) + I_n(t) \quad (4.2)$$

In this equation,  $I(t)$  is the stable house current after appliance switching event.  $\sum_{k=1}^M I_k(t)$  is the stable house current before the appliances were switched, and it is a mix current consisting of  $M$  appliances' working currents, which is represented by  $I_b(t)$ .  $I_{M+1}(t)$  is the working current of switched appliances and also the current that causes the house current  $\sum_{k=1}^M I_k(t)$  to change to a new state,  $I_{M+1}(t)$  is represented by  $I_a(t)$ .  $I_n(t)$  represents the system noise.

According to this equation (4.2), all appliances' working currents can be divided into two parts: the combined current in the circuit before an appliance is switched; the working currents of the switched appliances. Equation (4.2) thus lays the foundation for further switching event detection and appliance working current separation.

### **4.3 Appliance Switching Event Detecting Based on Heuristic Detection**

#### **Method**

The house current is represented as  $I(t)$  in Section 4.2.  $I(t)$  may fluctuate because of a switching event. The value of  $I(t)$  can be obtained by sampling at a constant frequency. As explained in the review on the principle of Heuristic detection in Chapter 2, each sampled value of  $I(t)$  in an observation period is first examined to find the maximum and minimum sampled data, and then the difference between these values is calculated to detect the switching event. During the detection process, it requires an auxiliary time meter and most of the data are duplicates. In order to improve the efficiency of this process and eliminate the use of the auxiliary time meter, this thesis employs the calculation of the house current rms value in each observation period to replace checking of each sampled data, and calculating the number of sample times can substitute for the auxiliary time meter.

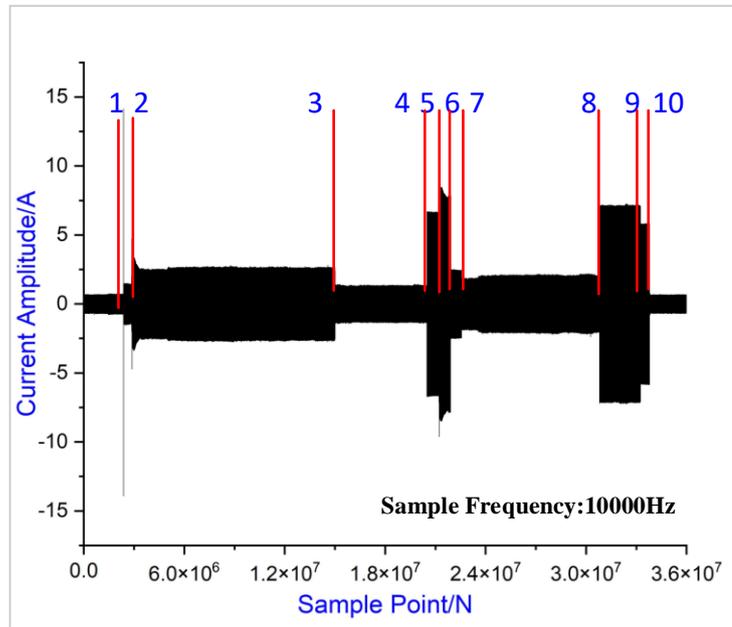
As for the detection process, the observation period is set as 2 seconds. Since the frequency is kept constant, the number of sampled data in each observation period is fixed, and the length of the observation period can then be determined by calculating the number of sample times. Based on the number and value of sampled current data in each observation period, the house current rms value  $i_{rms}$  is calculated using equation (3.1). After calculating the  $i_{rms}$  value within each observation period, the difference between  $i_{rms}$  values in adjacent periods can be calculated using:

$$\Delta i_{rms}(j) = i_{rms}(j) - i_{rms}(j - 1) \quad (4.3)$$

In equation (4.3),  $\Delta i_{rms}(j)$  is the difference value of  $i_{rms}$  between  $j$ th and  $(j - 1)$ -th observation periods.  $i_{rms}(j)$  and  $i_{rms}(j - 1)$  are the current rms value for  $j$ th and  $(j - 1)$ -th observation period respectively. When the  $i_{rms}$  value in  $j$ th and  $(j - 1)$ -th observation periods remains unchanged,  $\Delta i_{rms}(j)$  is zero, there is no switching event in  $j$ th observation period. By contrast if two adjacent results of  $i_{rms}$  values are different,  $\Delta i_{rms}(j)$  is non-zero, the switching event occurs at the  $j$ th observation period, and the sample point being with the max sampled data in  $j$ th period is detected as the switching point. The switching event is detected though comparing the  $i_{rms}$  difference, which replaces the checking of each sampled current data.

However, according to equation (4.2), both system noise and switching events can affect the house current. Because of such noise existence, the  $\Delta i_{rms}(j)$  between  $j$ th and  $(j - 1)$ -th observation period may never be zero. According to the original Heuristic principle, a threshold  $\varepsilon$ , commonly ranging from 0.1 to 1, is used to overcome the influence of noise. This threshold  $\varepsilon$  is thus applied in the proposed method to check the  $\Delta i_{rms}(j)$ . If the  $\Delta i_{rms}(j)$  between  $j$ th and  $(j - 1)$ -th observation period is larger than the threshold  $\varepsilon$ ,  $|\Delta i_{rms}(j)| > \varepsilon$ , a switching event may occur. When  $\Delta i_{rms}(j) > 0$ , some of the appliances may be switched on while others remain switched off.

Fig.4.2 is an example of an actual house current sequence in one hour. The sequence is sampled at the fixed frequency of 10000Hz, and the current signal to noise ratio (SNR) is high. Ten switching events occurred within that hour, as marked with red lines. The switching event detection method is used to analyse this current sequence.



**Fig.4. 2 One hour's house current sequence**

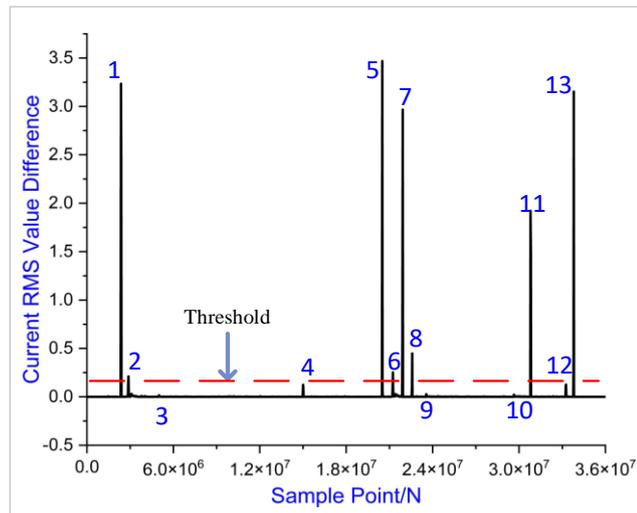
From Fig.4.2, the house current  $I(t)$  increases or decreases with the appliance switching events. Based on the sampled current data, the  $i_{rms}$  of each observation period is calculated, which includes 20000 sampled data. Table 4.1 compares the real switching event point and the detected point (with non-zero  $\Delta i_{rms}(j)$ ).

**Table 4. 1 Switching event detecting result**

Real Switching Event		Detected Switching Event	
Switching Event	No. of Switching Point	No. of point with non-zero $\Delta i_{rms}(j)$	The value of $\Delta i_{rms}(j)$
1	2702617	2710987	3.27
2	3029814	3037159	0.26
3	15281873	5029811	0.03
4	20873691	15291339	0.03
5	21022965	20881357	3.54
6	21738922	21032751	0.27
7	23281873	21747159	2.78
8	30962900	23290139	0.51
9	33541472	23929735	0.02
10	33762578	29189134	0.03
		30962900	1.73
		33541472	0.21
		33762578	3.23

In Table 4.1, all switching event points are obtained through that: the  $\Delta i_{rms}(j)$  between  $j$ th and  $(j - 1)$ -th period is calculated. If the  $\Delta i_{rms}(j)$  is non-zero, the sample point with the maximum sampled data in the  $j$ th period is detected as the switching point.

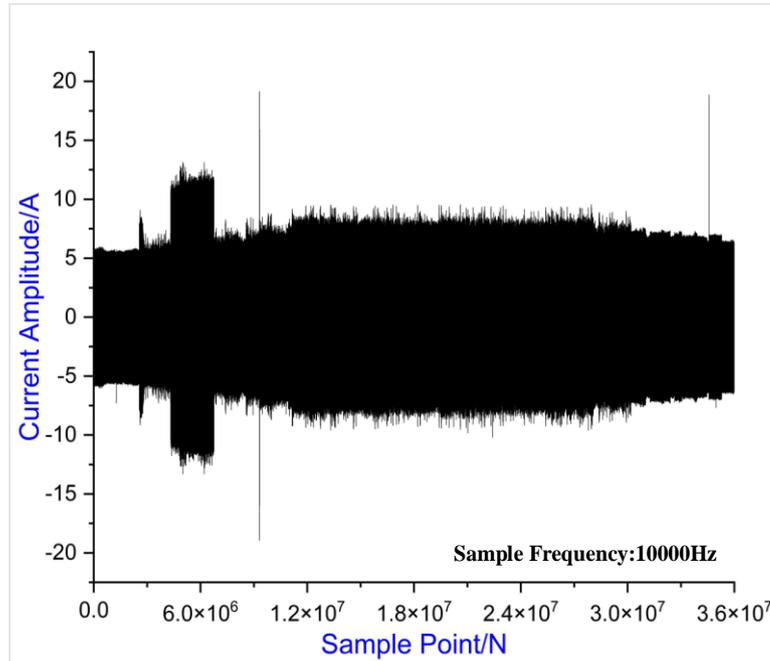
If the threshold was not applied, there would be 13 points detected as the switching points, when only 10 switching events happened. To prevent such noise fluctuation from being detected as an appliance switching event, the threshold was set as 0.1 to minimize this. Fig.4.3 compares all  $\Delta i_{rms}(j)$  values with the value of 0.1, the abscissa consists of all the sample points. If a sample point was detected as a switching point, the value of  $\Delta i_{rms}(j)$  is assigned to the corresponding ordinate value, otherwise, the ordinate value is 0. In Fig.4.3, with the threshold included, if the  $\Delta i_{rms}(j)$  are less than 0.1, the corresponding detected switching points would be screened out. The retained detecting switching points will match real switching event points, with a detection accuracy of 100%. The interval between the real switching event point and the detected switching point does not exceed 10000 sampling points, in other words, the time difference between the real switching event point and the detected switching point is less than one second. Hence, setting the threshold is an effective approach for detecting the switching event in a low noise system.



**Fig.4. 3 The  $\Delta i_{rms}(j)$  values compare with the threshold of 0.1**

The current sequence in Fig.4.2 was collected in a low noise environment with a high SNR. But for houses that are close to the harmonic source, the SNR can be low.

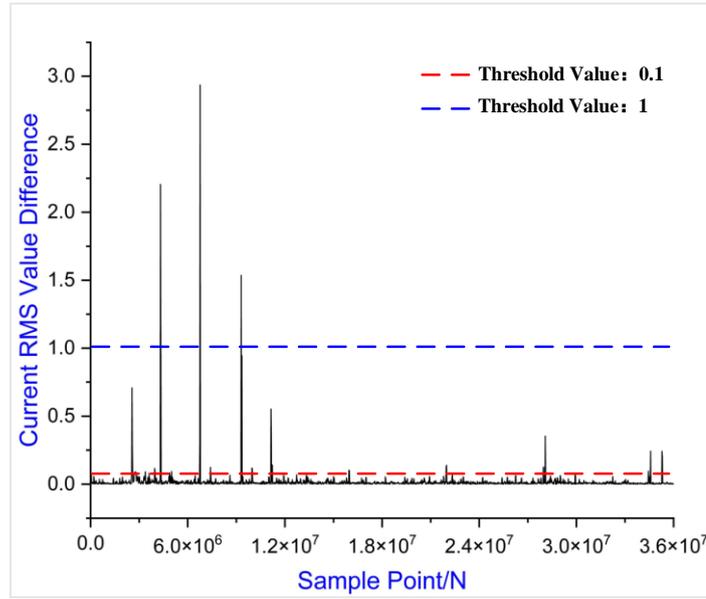
Fig.4.4 shows the current sequence sampled from a high noise power system for an hour. Though only eight switching events occurred during that hour, the house current was always fluctuating.



**Fig.4. 4 One hour’s current sequence with low SNR**

The sampled current sequence in Fig.4.4 is also processed by calculating  $i_{rms}$  values and checking  $\Delta i_{rms}(j)$  values to detect switching events. The observation period was also 2 seconds with 20000 sample points. When the  $i_{rms}$  between  $j$ th and  $(j - 1)$  - th adjacent observation periods is calculated, the  $\Delta i_{rms}(j)$  is always non-zero between two observation periods, almost each observation period contains a sample point that is detected as a switching point.

Fig.4.5 shows all sample points taken to each representing a switching event. The abscissa consists of all sample points, if a sample point was detected as a switching point, the value of  $\Delta i_{rms}(j)$  is assigned to the corresponding ordinate value, otherwise, the ordinate value is 0. Two threshold values are set as 0.1 and 1, which are marked as red and blue lines respectively in Fig.4.5. Two threshold values are used to screen out detected points caused by noise, completely different results were obtained under different thresholds.



**Fig.4. 5 The switching detection result under 0.1 and 1 threshold for current sequence with low SNR**

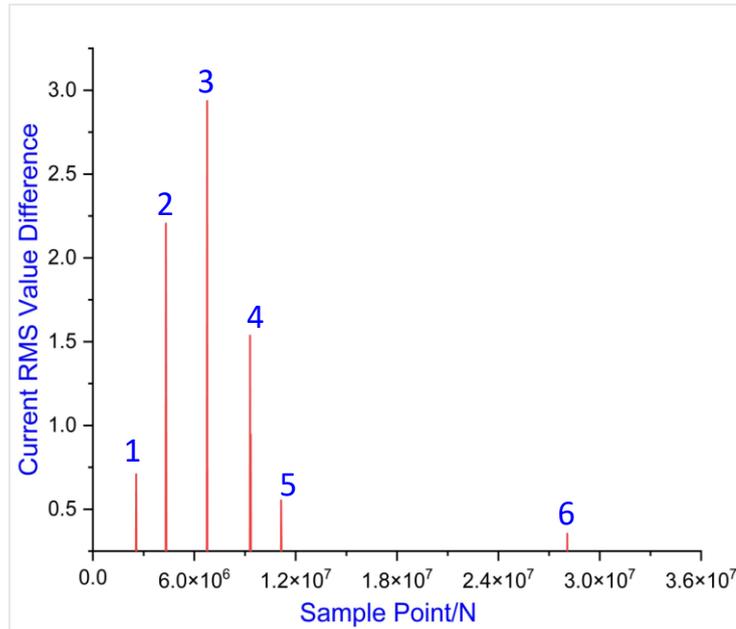
From Fig.4.5, there are many  $\Delta i_{rms}(j)$  between two observation periods above the threshold of 0.1. The noise influence and harmonic intrusion were detected as switching events. Although all the real switching events were detected, real events only account for 2% of the total detected points and 98% of the detected points had nothing to do with the switching events. These detected points were not only irrelevant but may also decrease the efficiency of detection. Conversely, in Fig.4.5, some of the  $\Delta i_{rms}(j)$  that are below the threshold of 1 may be missed. These are the switching events for small appliance working currents. Using this threshold, only 70% of the real switching events were detected, which is considered low detection accuracy. In summary, a small threshold value decreases the sensitivity of detection, while a large threshold value decreases the accuracy of detection. Detecting switching events and avoiding noise interference through a fixed threshold  $\varepsilon$  setting can be challenging in a high noise environment. It is often necessary to adjust the threshold according to the different system noise conditions. In practical applications, once the threshold value is set, it cannot be arbitrarily changed.

Besides using the threshold approach, the  $|\Delta i_{rms}(j)|$  between  $j$ th and  $(j - 1)$ -th adjacent observation periods can be compared with  $|i_{rms}(j)|$  of the first observation period and the  $|i_{rms}(j - 1)|$  of the last observation period. When  $|\Delta i_{rms}(j)|$  is larger

than one fifth of  $|i_{rms}(j)|$  or  $|i_{rms}(j-1)|$ , a switching event may occur. The switching event in high noise environment can thus be detected with a fixed threshold using the following equation:

$$|\Delta i_{rms}(j)| > \frac{1}{5} * |i_{rms}(j)| \cup |\Delta i_{rms}(j)| > \frac{1}{5} * |i_{rms}(j-1)| \quad (4.4)$$

Using a threshold value of 0.1 and this equation (4.4), the current sequence in Fig.4.4 was processed and analysed. The detected result is shown in Fig.4.6,



**Fig.4. 6 The switching detection result under threshold 0.1 and additional  $\Delta i_{rms}(j)$  comparing**

In Fig.4.6, where the  $\Delta i_{rms}(j)$  between two adjacent observation periods accords with equation (4.4), the sample point, which is the maximum sampled data in the latter period, will be detected as the switching point. The value of  $\Delta i_{rms}(j)$  is assigned to the corresponding ordinate value, otherwise the value is 0. Six points were detected and retained, where the detected points all represent an appliance switching event. Table 4.2 compares the real switching event point and the detected switching point. The time difference between the real switching event point and the detected switching point was less than a second. Although this was a high-noise environment, the detection accuracy and detection efficiency were satisfactory.

**Table 4. 2 The real switching event point and the detected switching point**

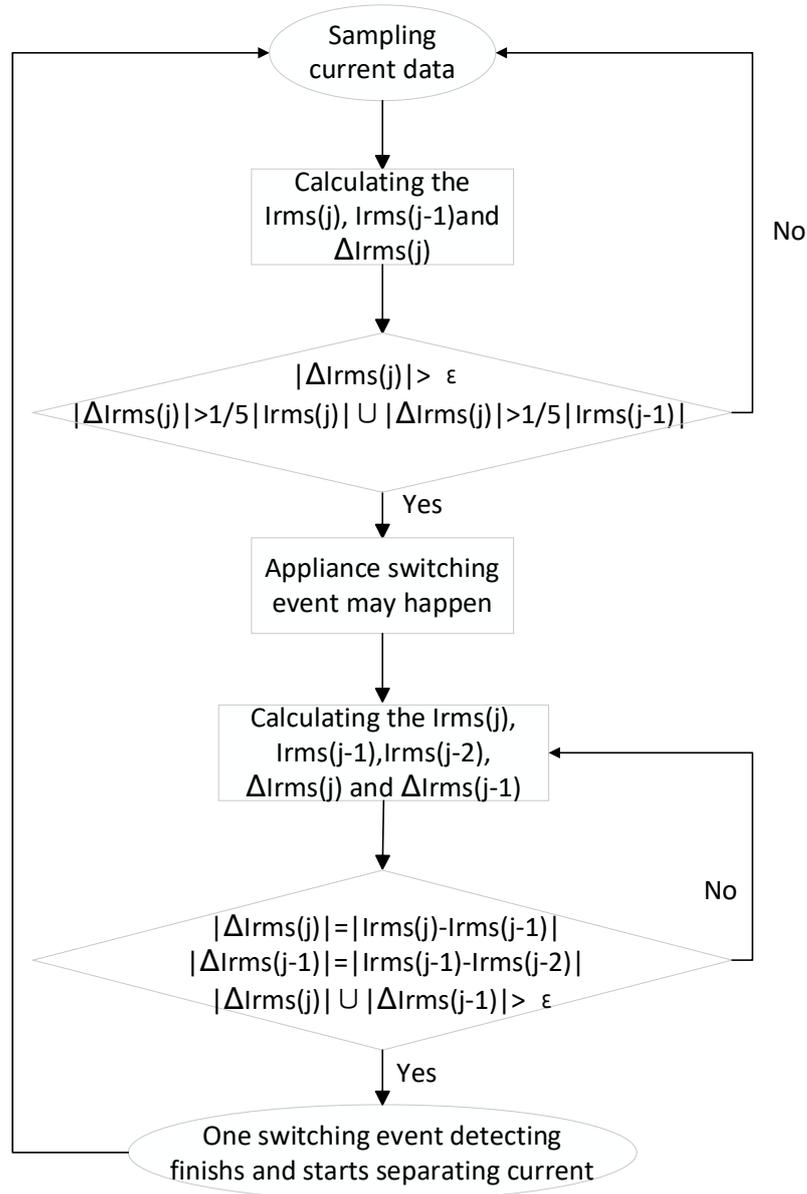
Real Switching Event		Detected Switching Event	
Switching Event	No. of data Point	No. of point with non-zero $\Delta i_{rms}(j)$	The value of $\Delta i_{rms}(j)$
1	2567152	2571275	0.71
2	4323874	4330162	2.26
3	6769432	6775793	2.93
4	9309814	9314821	1.54
5	11142372	11151618	0.55
6	28089843	28097987	0.35

After each switching event, the house current will undergo a sudden change and then quickly settle into its steady state. This process of switching to a steady state is known as the transient state. Once the transient state subsides, the house current will remain relatively stable, and detection of the next switching event begins. Therefore, determining whether a transient state is over or not becomes vital to the switching event detection process. The primary parameter for determining this is the  $|\Delta i_{rms}(j)|$  between two observation periods. When  $\Delta i_{rms}(j)$  and  $\Delta i_{rms}(j - 1)$  remains unchanged, the transient state is over. The switched appliance is in a steady state, and the house current should remain stable until the next switching event. To account for system noise, a determining threshold  $\gamma$  is used, as shown in equation (4.5):

$$\begin{aligned}
& |\Delta i_{rms}(j)| \cup |\Delta i_{rms}(j - 1)| < \varepsilon \\
& |\Delta i_{rms}(j)| = |i_{rms}(j) - i_{rms}(j - 1)| \\
& |\Delta i_{rms}(j - 1)| = |i_{rms}(j - 1) - i_{rms}(j - 2)|
\end{aligned} \tag{4.5}$$

The flow chart in Fig.4.7 provides an overall description of the proposed switching event detection method using the house current rms value. In Fig.4.7, the observation period is set as 2 seconds have been measured with the number of sample times, the house current rms value of each observation period is calculated, and then the rms value difference for the two periods is obtained. When the difference value meets the conditions of the threshold value and equation (4.4), an appliance switching event is considered to have occurred. After a switching event is detected, the house

current enters a transient state. The house current rms value in each observation period is calculated continuously, the rms value meets the condition in equation (4.5). When the house current goes into a steady state, the switching event detection is considered over.

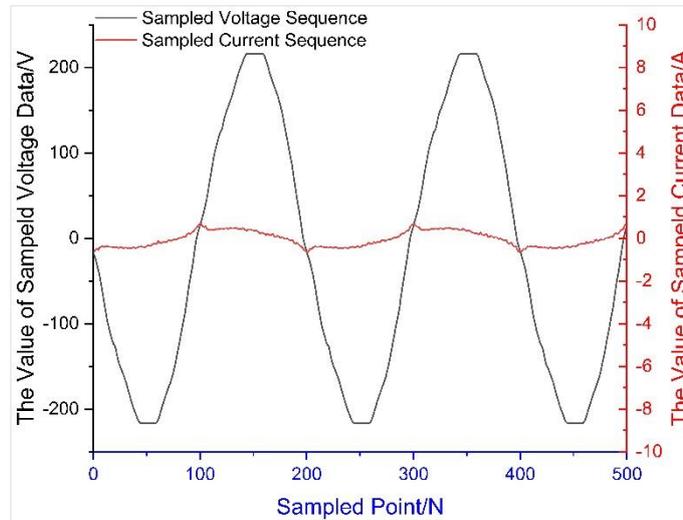


**Fig.4. 7 Flow chart of switching event detection**

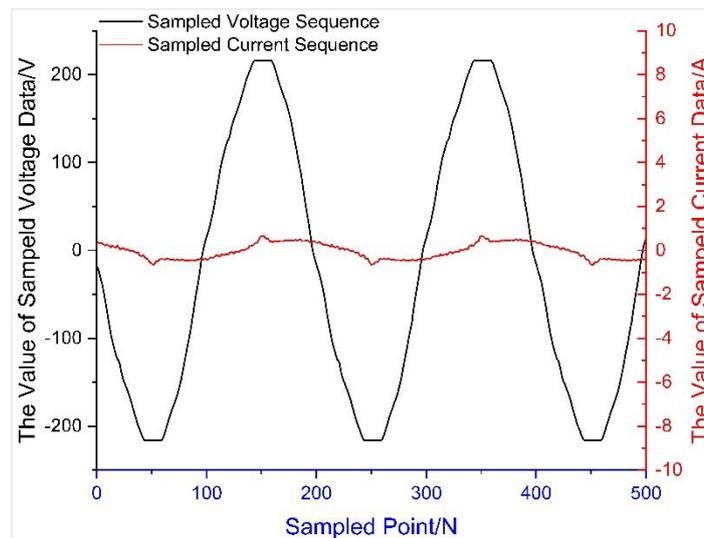
#### **4.4 Separating the Appliance Working Current through the Difference Method**

After detecting an appliance switching event, the working current of the switched appliance will then be separated to identify the appliance type. Using equations (2.7)

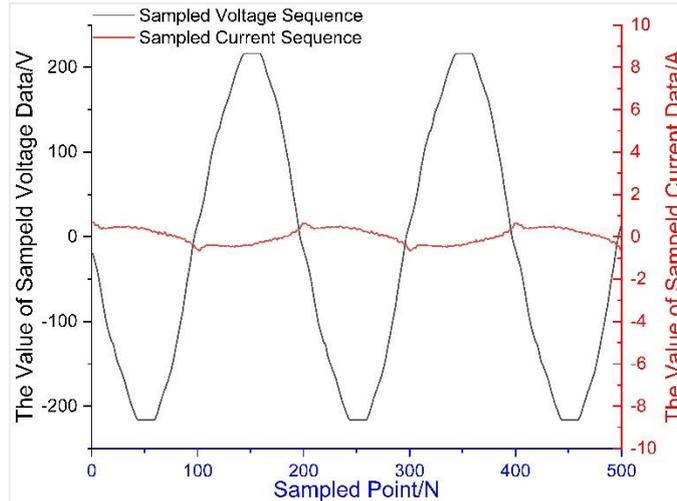
and (4.2), the working currents of switched appliances can be separated, in which the steady state house current sequence before switching is subtracted from the steady state current sequence after switching. In a residential power environment, all voltage and current signals are single phase sine waves. Fig.4.8, Fig.4.9 and Fig.4.10 show three sampled current and voltage signals.



**Fig.4. 8 The  $-90^\circ$  difference between current and voltage sequence phase angle**



**Fig.4. 9 The  $0^\circ$  difference between current and voltage sequence phase angle**



**Fig.4. 10 The 90° difference between current and voltage sequence phase angle**

In Fig.4.8, Fig.4.9 and Fig.4.10, the amplitude and frequency of the current and its voltage are fixed. However, as the switching of appliances is random, the beginning of each sequence sampling is also random. With sampling taken at different times, the phase angles of the sampled sequences can be different. The phase angle differences between these sampled voltage and current sequences are  $-90^\circ$ ,  $0^\circ$  and  $90^\circ$  respectively. Although all these sequences were sampled from the same current and voltage signals, the features of the obtained sequence vary due to the different sampling times. This reduced the accuracy for further identification, so separating the appliance working current under varying phase angle is thus very challenging. In order to avoid the features distortion caused in the sampling process, the voltage signal is used as the reference.

As for residential power systems, the voltage parameter is fixed at standard values such as 230V/50Hz in the UK and 120V/60Hz in the US. Hence, the house voltage  $V(t)$  remains stable regardless of the number of switching events. If the voltage sequence is sampled from the same voltage value, the phase angle of the sampled sequence will stay unchanged. In this thesis, the voltage sequence is acquired by sampling starting from a zero voltage and in an increasing trend. The corresponding conditions are as follows:

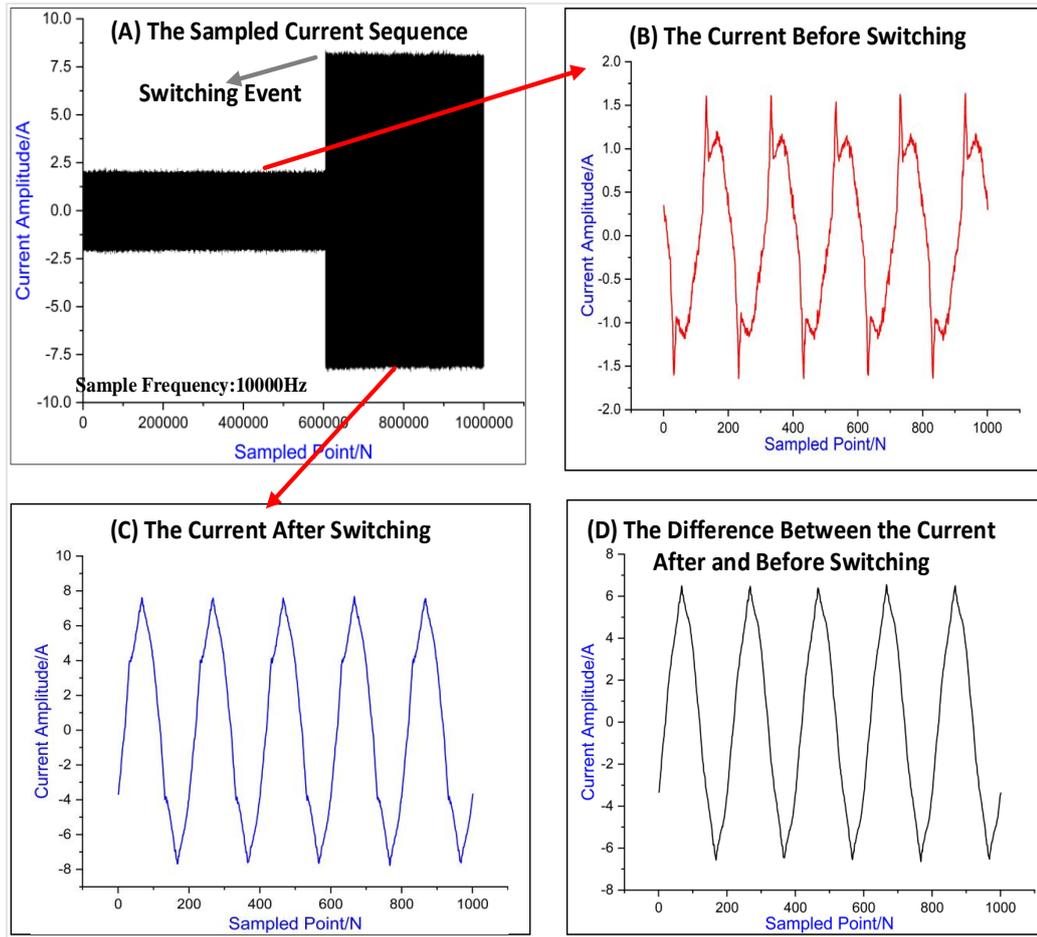
$$\begin{aligned}
V(t - 1) &< 0 \\
V(t) &= 0 \\
V(t + 1) &> 0
\end{aligned} \tag{4.6}$$

When the voltage value meets the conditions in (4.6),  $V(t)$  will be sampled. The sampled voltage sequence is  $V = \{V_1, V_2, V_3, \dots, V_L\}$ , where  $L$  is the number of sample times, and  $V_1$  to  $V_L$  are the sampled voltage values. All these data compose a periodic voltage sequence with a phase angle of  $0^\circ$ . Therefore, the phase angle of sampled house voltage sequence is  $0^\circ$ , if the sampling of the house voltage and current start at the same time, with the phase angle of the sampled voltage sequence fixed at  $0^\circ$ , the phase angle of the sampled house current sequence can then be determined accurately. A current sequence  $I = \{I_1, I_2, I_3, \dots, I_L\}$  can be obtained with its phase angle, where  $L$  represents the number of data points in the sequence. Using this approach, the problem of the uncertain phase angle of the house current in the sampling process is resolved. When the house current sequence can then be sampled with the fixed phase angles, the working current of switched appliance can now be separated.

According to equation (4.2),  $I_b(t)$  is the stable house current before appliances switching,  $I_a(t)$  is the stable house current after appliances switching. Based on the sampled sequence of  $I_a(t)$  and  $I_b(t)$ , The working current sequence of switched appliances can be separated according to equation (2.7), which is the difference between the steady house current before and after switching, expressed as follows:

$$[I_a] - [I_b] = [I_a^1 - I_b^1, I_a^2 - I_b^2, \dots, I_a^L - I_b^L] \tag{4.7}$$

In equation (4.7),  $[I_a] = [I_a^1, I_a^2, \dots, I_a^L]$ , it represents the sampled steady state current sequence after the appliance was switched.  $[I_b] = [I_b^1, I_b^2, \dots, I_b^L]$ , it represents the sampled steady state current sequence before the switching event.  $[I_a^1 - I_b^1, I_a^2 - I_b^2, \dots, I_a^L - I_b^L]$  is the working current sequence of the switched appliances, which causes the current sequence  $[I_b]$  changing to the current sequence  $[I_a]$ . Fig.4.11 shows a current separation example, the house current sequence was processed by using this sampling approach to obtain the working currents of the switched appliances, the length of the current sequence is 100 seconds.

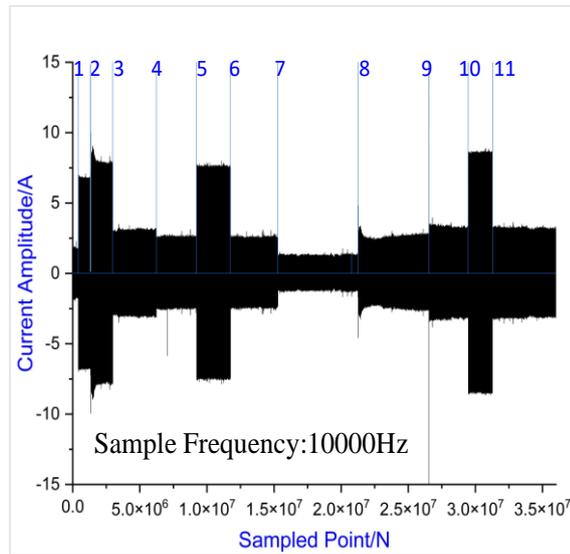


**Fig.4. 11 Appliance working current separating process**

In Fig.4.11A, the current sequence includes 1 million sample points, which are sampled at 10000Hz (100 seconds), and the switching event happened in the 601479<sup>th</sup> sampled point. In Fig.4.11B, the sequence includes 1000 points (0.1 seconds), the sampled sequence is the house current before the switching event, which is also the working current of a single appliance. In Fig.4.11C, the sequence includes 1000 points, the sampled sequence is the house current after the switching event. It is also the sum of the switched appliance working currents and the appliance working currents before the switching event. In Fig.4.11D, the sequence includes 1000 points. The sequence is the difference between the current before and after the switching event, which would be the working current of the switched appliance. The obtained sequences had regular and periodic waveforms, indicating that the separated sequences are not noise but the actual currents of the appliances.

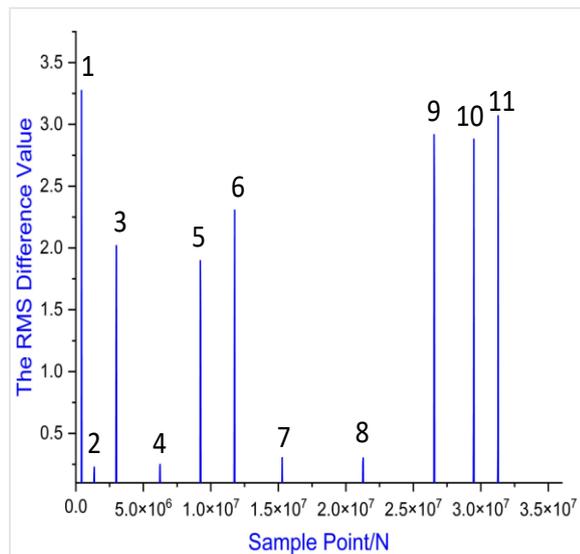
#### 4.5 Testing of Switching Event Detection and Working Current Separation

This section presents the testing results of the detecting and separating process using the UK-DALE [57] dataset. Fig.4.12 shows the house current waveform within an hour from 18:00-18:59, Each switching event is marked with a blue line.



**Fig.4. 12** The one hour's current sequence

Using the sequence in Fig.4.12, the proposed method first detects the switching event, the detecting results are shown in Fig.4.13,



**Fig.4. 13** The switching event detecting result

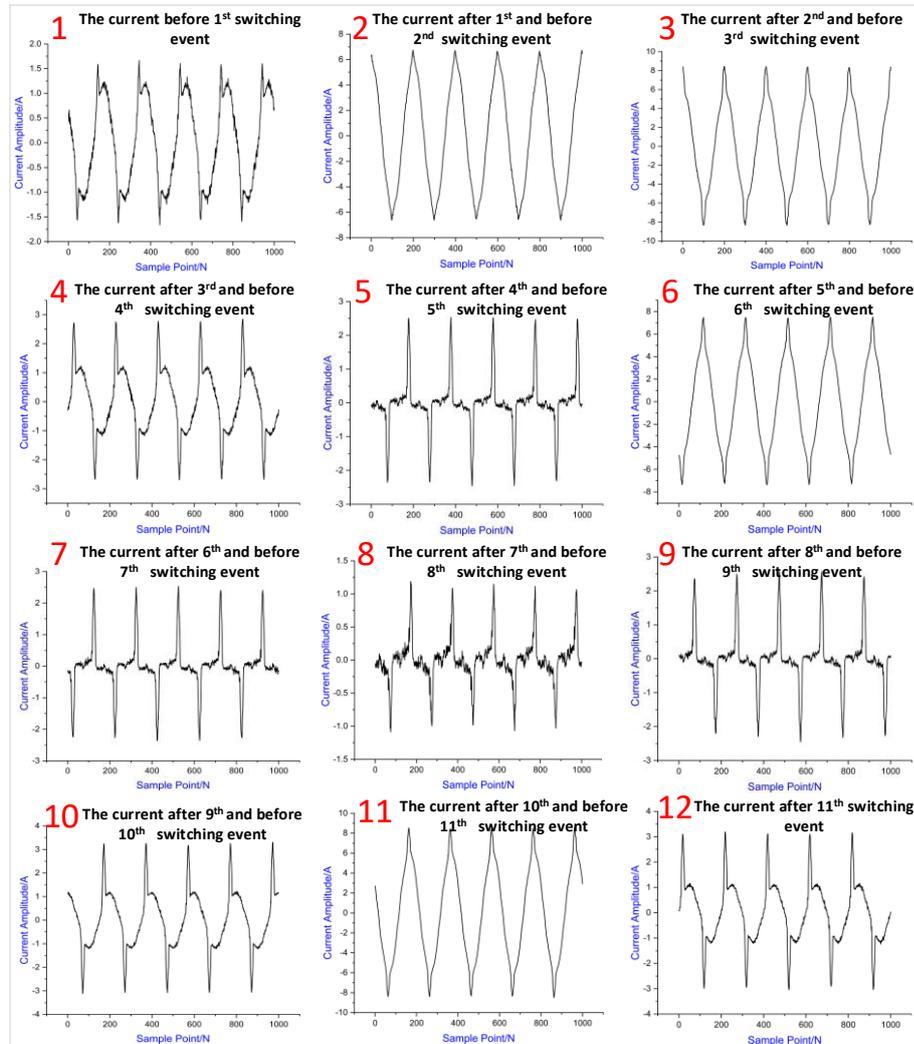
In detection process, the observation period is 2 seconds with 20000 sample points. When the  $i_{rms}$  between two adjacent observation periods is obtained, the  $\Delta i_{rms}(j)$  between  $j$ th and  $(j - 1)$ -th adjacent period is calculated. If the  $\Delta i_{rms}(j)$  is over the threshold value of 0.1 and accords with equation (4.4), the sample point with the maximum sampled data in the latter period is detected as the switching point. In Fig.4.13, the abscissa consists of the sample points, and if a sample point is detected as a switching point, the corresponding ordinate value of  $\Delta i_{rms}(j)$  is assigned, otherwise given a value of 0. Table 4.3 compares the real switching event point in one hour and the point been with non-zero  $\Delta i_{rms}(j)$ .

**Table 4. 3 The real switching point and the detected switching point**

Real Switching Event		Detected Switching Event	
Switching Event	No. of data Point	No. of point with non-zero $\Delta i_{rms}(j)$	The value of $\Delta i_{rms}(j)$
1	4200987	4213687	3.27
2	1346350	1352645	0.26
3	2980401	2991389	2.03
4	6220782	6234571	0.31
5	9220600	9223148	1.95
6	11741369	11753722	2.19
7	15280251	15289784	0.33
8	21263691	21274781	0.35
9	26540028	26551195	2.87
10	29461139	29469852	2.71
11	31285391	31289171	3.12

The detected switching points all match real switching event points, the detection accuracy is 100%. Furthermore, the interval between the real switching event point and the detected switching point does not exceed 20000 sampling points, so the time difference between the real switching event point and the detected switching point is less than two seconds.

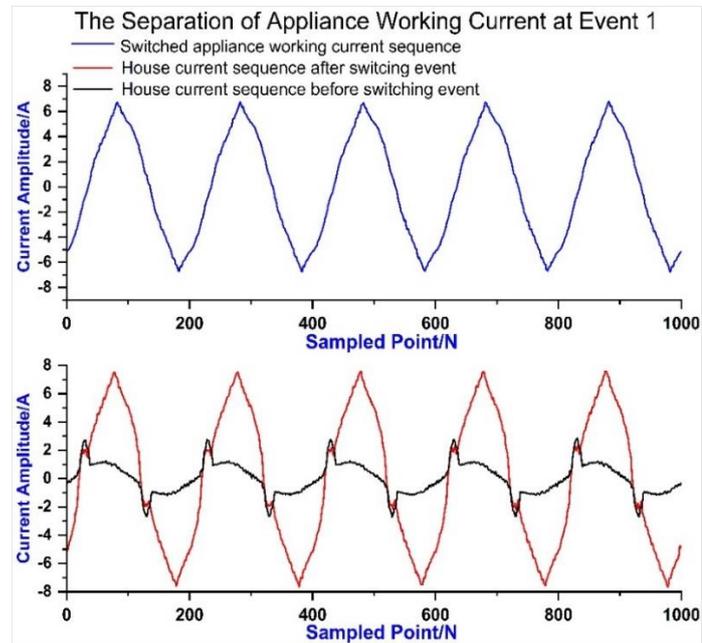
Once the switching event is detected, the switched appliance working current will be sampled according to equations (4.6) and (4.7). Fig.4.14 shows 12 sampled current sequences which were taken before and after the 11 switching events.



**Fig.4. 14 Working current sequence before and after 11 switching events**

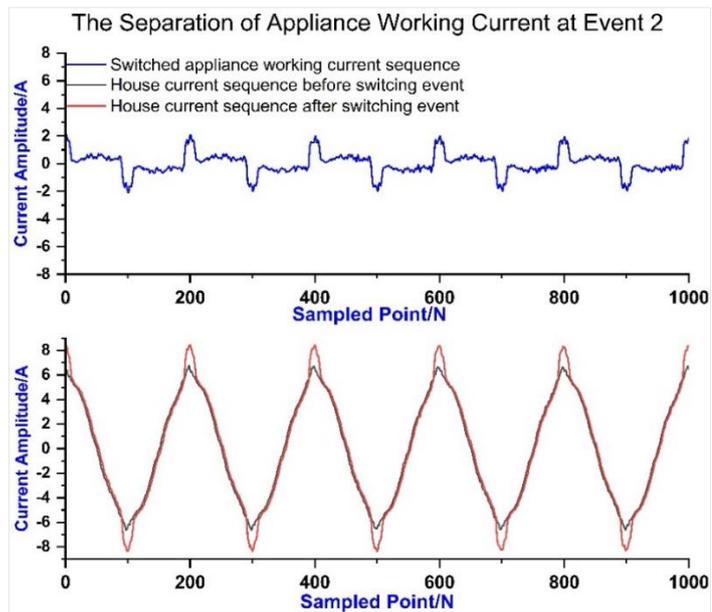
Using equation (4.8), the working current of switched appliances is separated based on the current sequences before and after each switching event. The steady state house current sequence before the switching event is subtracted from the steady state current sequence after the switching event. Fig.4. 15, Fig. 4.16 and Fig.4.17 shows the current separation process after the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> switching events, with each one of them representing respectively the event of appliance switching ON, the appliance operation state changing and the appliance switching OFF.

In Fig.4.15, the black curve represents the steady state house current sequence before the switching event. When switching event 1 occurred and was detected, the house current was then sampled, and the sampled current sequence showed as the red curve. The current sequence before switching was subtracted from the current sequence after switching, the subtracting result is shown as the blue curve, which represents the switched appliance working current. The steady state current after the switching event is then stored for the next separation process.



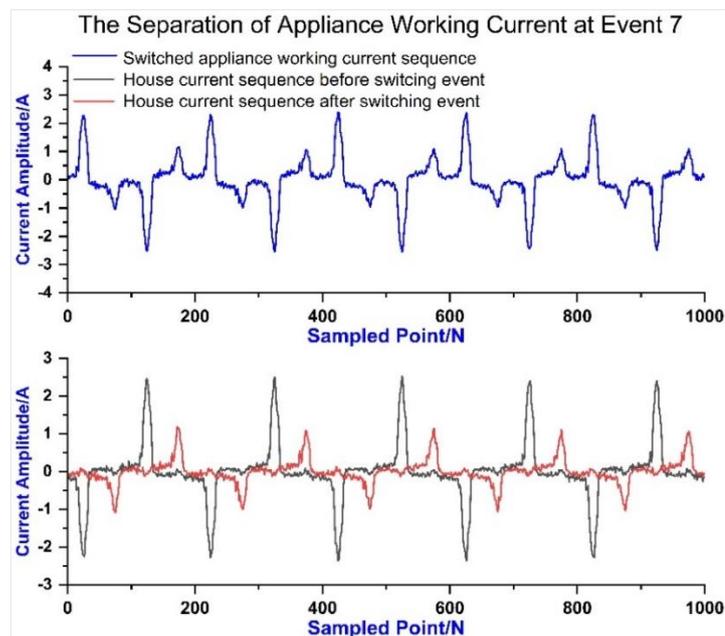
**Fig.4. 15 The current separation after 1<sup>st</sup> switching event**

In Fig.4.16, the switching event was just a change in the working state of the appliance, and hence the current sequences before and after the event were almost identical except for their amplitudes. The black curve is the steady state house current sequence before the appliance operation state change, the red curve is the steady state house current sequence after the appliance operation state change. The current sequence before running state change also was subtracted from the current sequence after running state change, using the same separation process as the switched appliance working current. The resulting separated sequence was periodic, with no significant distortion, allowing for the separation of some small current changes using the difference method.



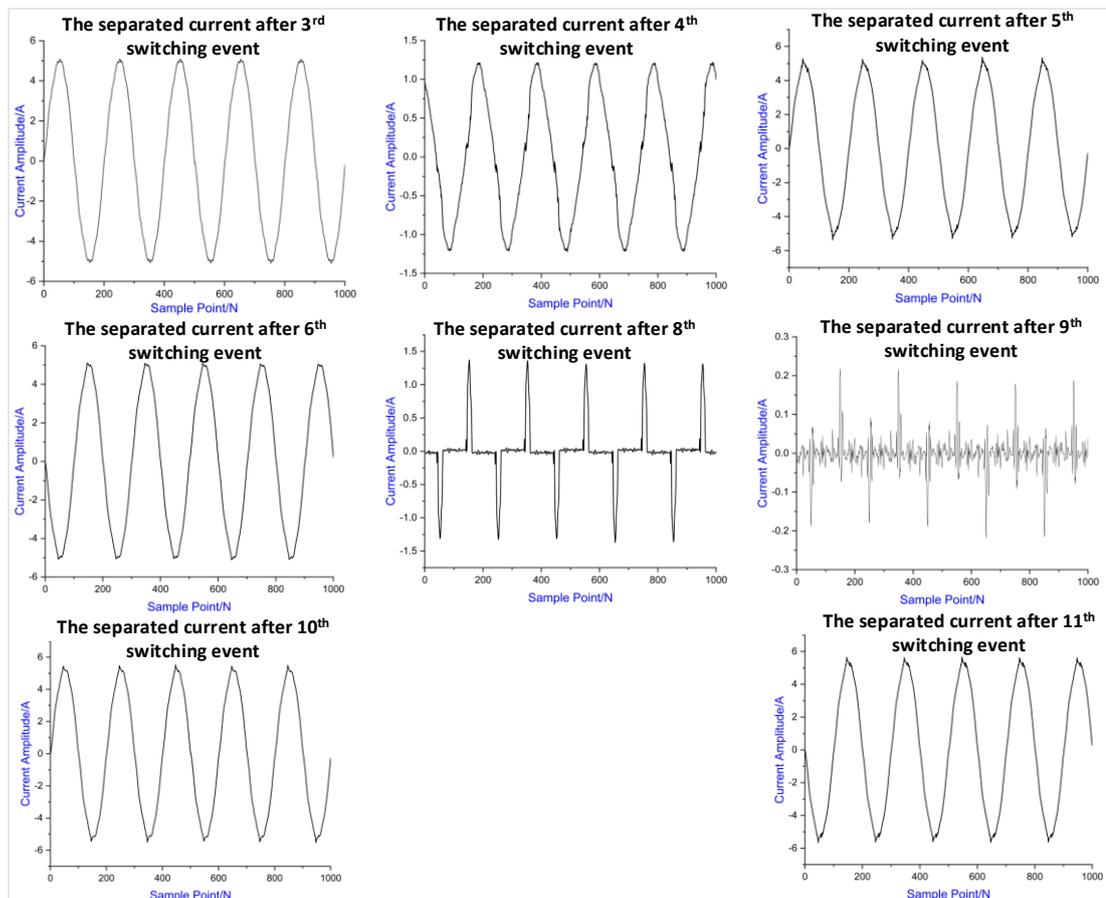
**Fig.4. 16 The current separation after 2<sup>nd</sup> switching event**

In Fig.4.17, the black curve represents the steady state house current sequence before the appliance switching OFF event. When the switching OFF event occurs, the current sequence before switching OFF was subtracted from the current sequence after switching OFF and the resulting difference gives the working current of the switched off appliance.



**Fig.4. 17 The current separation after 7<sup>th</sup> switching event**

The other separated current sequences are shown in Fig.4.18, and the current waveforms generally coincide with the sine waveform, The separated sequences are regular and periodic. However, there are obvious differences in the shapes of the current waveforms of some of the sequences, such as sequences 8 and 9. On the other hand, other sequences are almost identical, making it difficult to distinguish them directly. Therefore, the separated waveforms need to be further analysed for the identification purposes.



**Fig.4. 18 The separated current sequences after switching event**

#### 4.6 Summary

This chapter presents the detection and separation of switched appliance working current. Firstly, the original Heuristic method is improved to achieve a high detection sensitivity and less detection response time. These improvements are explained and tested using real residential electric data. Secondly, when an appliance is switched and

the switching is detected, the house current is different before and after switching, so subtracting the house current before a switching event from the house current after the switching event, provides the working current of the switched appliance. The entire separation process is based on the difference method, in which the shortcoming of not being able to obtain the phase angle of separated current is resolved. In the next chapter, the separated appliance working current feature will be identified.

## Chapter 5 The Switched Appliance Type Identification

### 5.1 Introduction

With the working current of a switched appliance separated, the frequency feature of the separated current can thus be used to identify the type of appliance. However, current frequency features may overlap between some appliances. To improve identification accuracy, other than the appliance's current frequency features, the user behaviour and appliance operation pattern will also be incorporated into the identification process. There are three parts to this chapter. Firstly, the usage behaviour and running pattern of household appliances are analysed, providing the foundation for further appliance operation pattern regressing and identification modification. Secondly, by using the appliance working current frequency, the K-nearest neighbor (KNN) method can be used to achieve basic identification of appliance types. The KNN model is trained and reconstructed to adapt to the limited online storage space and reduce the response time of identification. Thirdly, regressing the appliance operation pattern with a back-propagation neural network (BPNN) to modify and correct the identification results.

For the first part, the non-electric data, from the public dataset UK-DALE [21], on appliance switching is used to analyse the usage habits of customers and the appliance running pattern. This analysis shows that the customer behaviour and appliance running pattern is an effective strategy for appliance identification. The specific analysis and process are discussed in Section 5.2.2.

For the second part, a training set is formed using the known appliance features, the KNN model is constructed and trained based on this. The KNN model compares the unknown appliance features with those of each known appliance and makes an identification based on the best match. The general process of KNN model building and appliance identification is given in more detail in Section 5.3.1, while this specific approach for appliance identification is tested in Section 5.3.2.

In the third part of this chapter, a BPNN is used to reproduce appliance operation patterns from the appliance switching data. The general process of building the BPNN

model and regressing the operation pattern curve are described in sections 5.4.1 and 5.4.2. The obtained operation pattern curve is validated and tested in Section 5.4.3. The basic identification result of the KNN method can then be corrected, to avoid mistakes caused by overlaps in the appliances' electric features. The modification and correction methods are tested in section 5.4.4.

## 5.2 Analyse the usage habits of customer and the appliance running pattern

UK-DALE dataset is the data resource for this analysis, and it contains the electric and non-electric data of appliances in five different households in London. Table 5.1 shows the data information used in this thesis.

**Table 5. 1 The UK-DALE data set details [21]**

	<b>House A</b>	<b>House B</b>
<b>Date of start measurement</b>	2012-11-09	2013-02-17
<b>Date of end measurement</b>	2015-01-05	2013-10-10
<b>Duration of measurement</b>	786	234
<b>Number of appliances</b>	53	19
<b>Electric parameters</b>	House current, House voltage, Total power, Appliance power	House current, House voltage, Total power
<b>Non-electric parameters</b>	Appliance switching time	Appliance switching time

Different switching regularities imply different usage behaviours. The appliance operation patterns and switching probabilities can be obtained by analyzing usage behaviours, as shown in Fig.5.1. It shows the statistical results of the switching times of four appliances, namely the oven in House A, the kettle in House A, the oven in House B and the kettle in House B. This provides a rough profile of the appliance operation patterns and the appliance switching probabilities of different periods, and it is a unique feature for the different appliances.

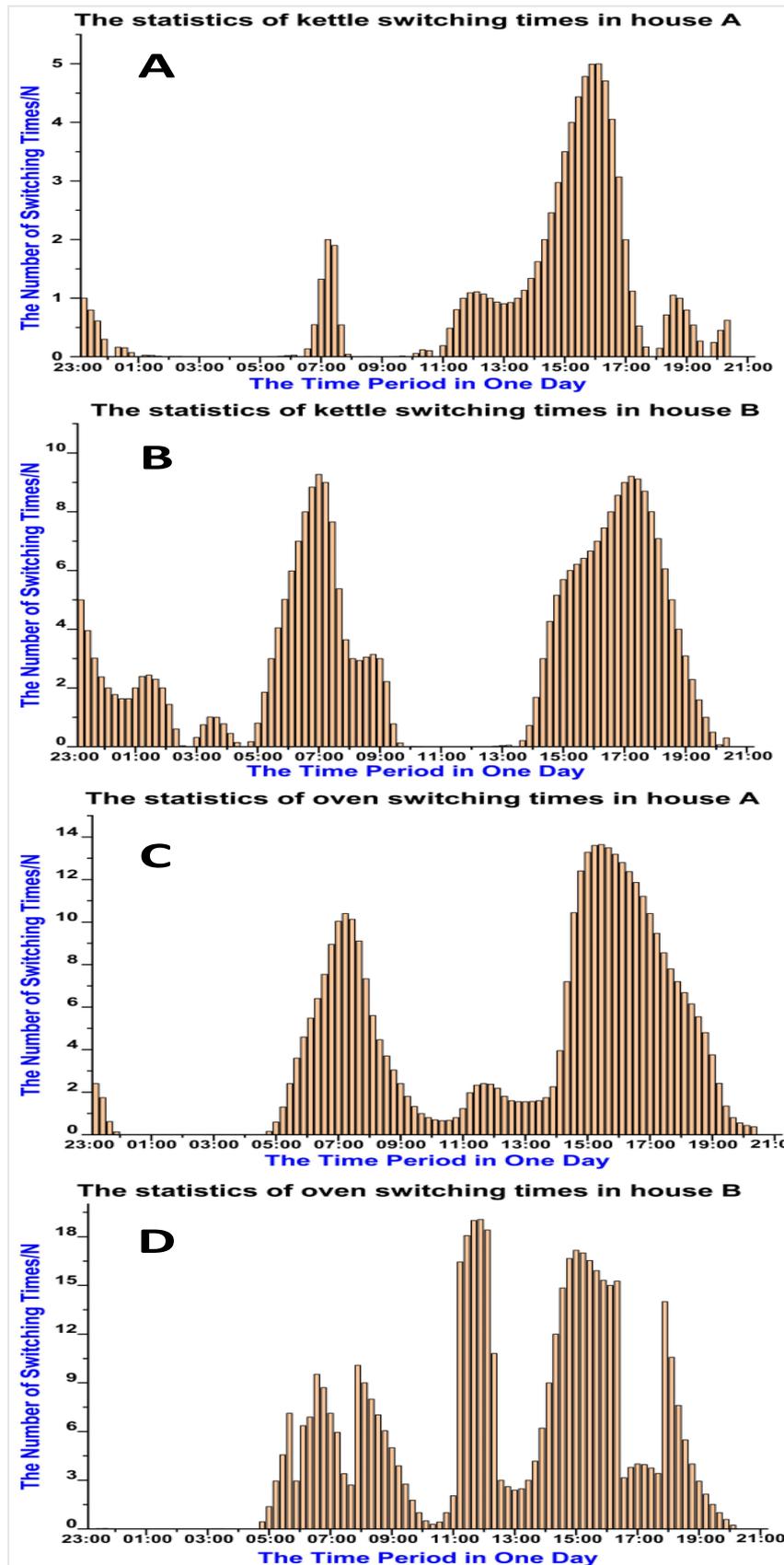
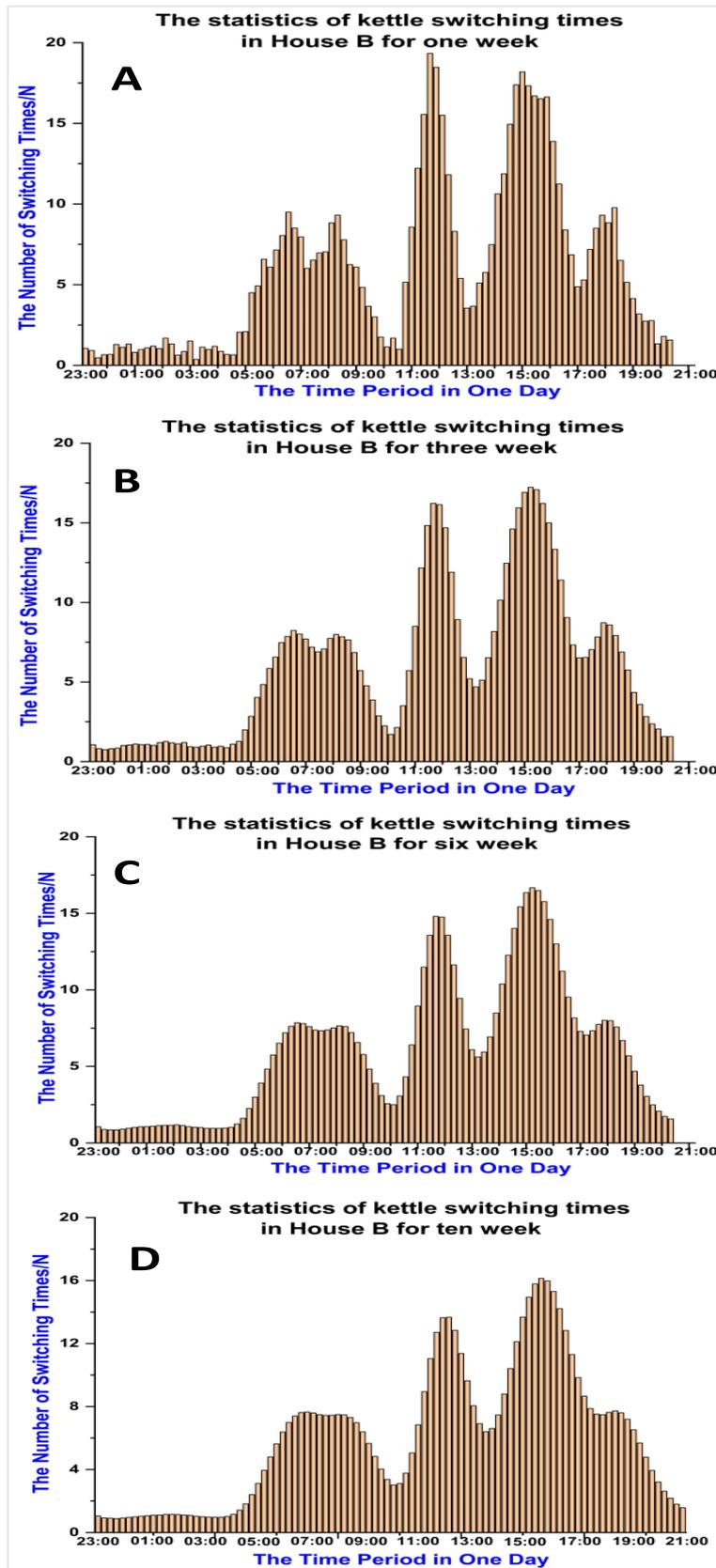


Fig.5. 1 Four appliances switching times in one week

In Fig.5.1, the switching times of the oven (Fig.5.1 C and Fig.5.1 D) and kettle (Fig.5.1 A and Fig.5.1 B) were counted for one week in two households. The ordinate represents the number of times of appliance switching, the abscissa is the time period of a day. From the corresponding axis, the histogram shows the switching times of the kettle and the oven for every quarter of an hour. These four diagrams show completely different trends and waveforms. Although these four appliances are similar in electric features, the switching times of these four appliances are different for each quarter-hour, due to different user behaviour and operation pattern. Furthermore, the appliance usage is random and switching events are not well distributed throughout the day. The switching of appliances is concentrated during specific time periods, for example, the switching of the oven in house A is mostly around 4 pm (Fig.5.1 A), and the switching of the oven in house B is around 5 pm (Fig.5.1 B), which are consistent with the living habits of the customers.

The appliance usage habits of the customer and a rough regress of the appliance operation pattern can be mined from the statistics of the appliances switching times to determine the periods when the switching probability of the appliances is higher. However, such simple statistics cannot reflect the switching regularity of all appliances. From Fig.5.1.D, the switching regularity of the appliance is lower compared to the other three, in other words, the operation pattern of the kettle in House B is not obvious. Without clear switching regularity of the appliance, it is challenging to regress an accurate operation pattern, and the calculation of switching probability based on this operation pattern becomes meaningless. Increasing the statistical period can effectively enhance the regularity of the appliance switching data for more accurate operation patterns and clearer customer usage habits. Fig.5.2 shows the statistical results of the kettle's switching times in House B for different periods: one week, three weeks, six weeks and ten weeks.



**Fig.5. 2** The statistical results of the kettle’s switching times for one week, three weeks, six weeks and ten weeks

In Fig.5.2 A, the switching times of the kettle in one week were counted, which did not produce significant patterns. By extending the statistical period, the user behaviour and appliance operation patterns were fully extracted, with the regularity becoming clearer moving from Fig.5.2.B to Fig.5.2.D. Increasing the statistical period enables some small-probability events of the appliance switching at certain time points to be included in the statistics, for example, there was an oven switching event at 4 am. With such inclusions, the obtained operation pattern becomes closer to the real running conditions. By considering the probability of appliances being switched at any time of the day, the calculation of the switching probability distribution (Fig.5.2.D) becomes more accurate. However, a period of ten successive weeks may be too long to be realized in practical applications. It is thus vital to use more intelligent algorithms to process shorter-term statistical data to regress accurate appliance operation patterns and switching probability distribution curves. The BPNN demonstrates a strong learning capability to do exactly this and hence will be applied in appliance operation pattern regressing for short-term statistical data.

### **5.3 Identification of Appliance Type**

The KNN method [96] is used to classify and identify the separated current, then the type of switched appliance is obtained. The structure of the KNN model is simplified, the learning strategy of the KNN model is redesigned, and the classifying rule of KNN is modified. With all improvements, the fast online identification is achieved, and the improved KNN method is more suitable for the residential application scenario.

#### ***5.3.1 The KNN Method reconstructing and separated current identifying***

The separated current is identified through the KNN method in these four steps. Firstly, the KNN model is trained and reconstructed by the training samples' features set. Secondly, the electric feature difference between the unknown sample (the separated current) and all training samples is measured. Thirdly, by setting a specific value  $K$ , the  $K$  training samples that are closest to matching the features of the unknown sample are found. Lastly, the unknown sample is classified into the feature category it most closely aligns with.

The first step of this process is to obtain currents from different known appliances and then expand them using the Fourier series. Table 5.2 lists eight appliances and their respective appliance harmonic components, which were analysed into their current sequences in Chapter 3. These currents are measured from real appliances in actual residential power environment. The first to fourth harmonic components of each appliance's working current is shown, and the unit of the harmonic component is in ampere.

**Table 5. 2 First to fourth harmonics of eight appliances**

<b>Appliance</b>	<b>1st Harmonic</b>	<b>2nd Harmonic</b>	<b>3rd Harmonic</b>	<b>4th Harmonic</b>
TV	0.345	0.009	0.310	0.008
Refrigerator	1.205	0.033	0.079	0.007
Microwave Oven	7.620	0.682	2.943	0.103
Air-conditioning	4.477	0.691	0.334	0.187
Laptop	2.769	0.172	0.432	0.272
Electric Kettle	5.146	0.005	0.089	0.001
Geysers	2.301	0.538	0.139	0.015
Vacuum Cleaner	6.513	0.728	2.138	0.236

In Table 5.2, each known appliance serves as a training sample in the KNN model. The features of the training samples are represented by the first to fourth harmonics. Once the feature of the training sample has been learned, the training and constructing of the KNN model is complete. The learning process involves using physical storage space to record each feature element value of each training sample. This thesis uses one training sample with four elements to minimize storage space and reduce response time for later steps.

In the second step, the feature difference between an unknown sample (the separated working current of switched appliance) and the training samples is measured using the Euclidean distance. The Euclidean distance between any appliance features can be calculated using equation (2.9). Table 5.3 compares the Euclidean distance between 8 training samples with each other.

**Table 5.3 The Euclidean distance between 8 training samples to each other**

	<b>TV</b>	<b>RE</b>	<b>MO</b>	<b>AC</b>	<b>LAP</b>	<b>GE</b>	<b>EK</b>	<b>VC</b>
<b>TV</b>	0	0.8603	7.3061	4.1879	2.4240	4.8011	6.2046	6.2056
<b>RE</b>	0.8603	0	6.4477	3.3375	1.5642	3.9410	5.3475	5.3486
<b>MO</b>	7.3061	6.4477	0	3.1430	4.8975	2.5577	1.1070	1.1070
<b>AC</b>	4.1879	3.3375	3.1430	0	1.8391	0.9384	2.0360	2.0360
<b>LAP</b>	2.4295	1.5702	4.8777	1.7851	0	2.3811	3.7786	3.7798
<b>GE</b>	4.8010	3.9411	2.5650	0.9582	2.3770	0	1.5255	1.5295
<b>EK</b>	6.2046	5.3475	1.1070	2.0360	3.7786	1.5255	0	0.8603
<b>VC</b>	6.2056	5.3486	1.1070	2.0360	3.7798	1.5295	0.8603	0

In Table 5.3, the Euclidean distance intuitively reflects the electric feature difference or similarity between appliances. By measuring the Euclidean distance between the unknown sample (the separated working current of switched appliance) and each training sample, it reveals the degree of feature matching between them. These measured results are used in the next step of the process.

In the third step, the value of K was set to 3 in the KNN model [97]. Based on the Euclidean distance calculated earlier, the first K training samples nearest to the unknown sample will be determined. If the nearest Euclidean distance exceeds a certain value, the unknown appliance will not be identified as it may belong to other appliance types not included in the training sample set.

In the fourth step, the unknown sample is identified to be a certain appliance category through majority voting [98] of K training samples. If the K samples belong to the same category, there is a high possibility that the unknown sample also belongs to this category. However, the number of home appliances is limited, and there may be only one or two appliances in each category, so the K training samples are always from different categories. The number of samples in each category may also be almost identical and hence decrease the reliability of majority voting.

In order to improve the reliability of majority voting and the accuracy of identification, different weighting is given to different training samples in the majority voting process. The Euclidean distances between the unknown sample and the K training samples are also not the same. The larger the Euclidean distance, the greater

the difference between the unknown sample and the training sample. This means the probability that the unknown sample and training sample belong to the same category is low. Hence, training samples with a larger Euclidean distance to the unknown sample carry a lower weighting in the majority voting process. Conversely, the training samples with small Euclidean distances to the unknown sample carry a higher weighting in the majority voting process. To obtain the weighting of each training sample for the majority voting process, all Euclidean distances need to be normalized. Equation (5.1) is used to achieve the normalization,

$$P_{x_i,m} = \frac{w(x_i,\psi_m)}{\sum_{m=1}^K w(x_i,\psi_m)} \quad (5.1)$$

where  $P_{x_i,m}$  is the normalization result,  $w(x_i,\psi_m)$  is the reciprocal of the Euclidean distance between the training sample  $\psi_m$  and the unknown sample  $x_i$ ,  $m = 1,2 \dots, K$ .  $w(x_i,\psi_m)$  can be calculated by

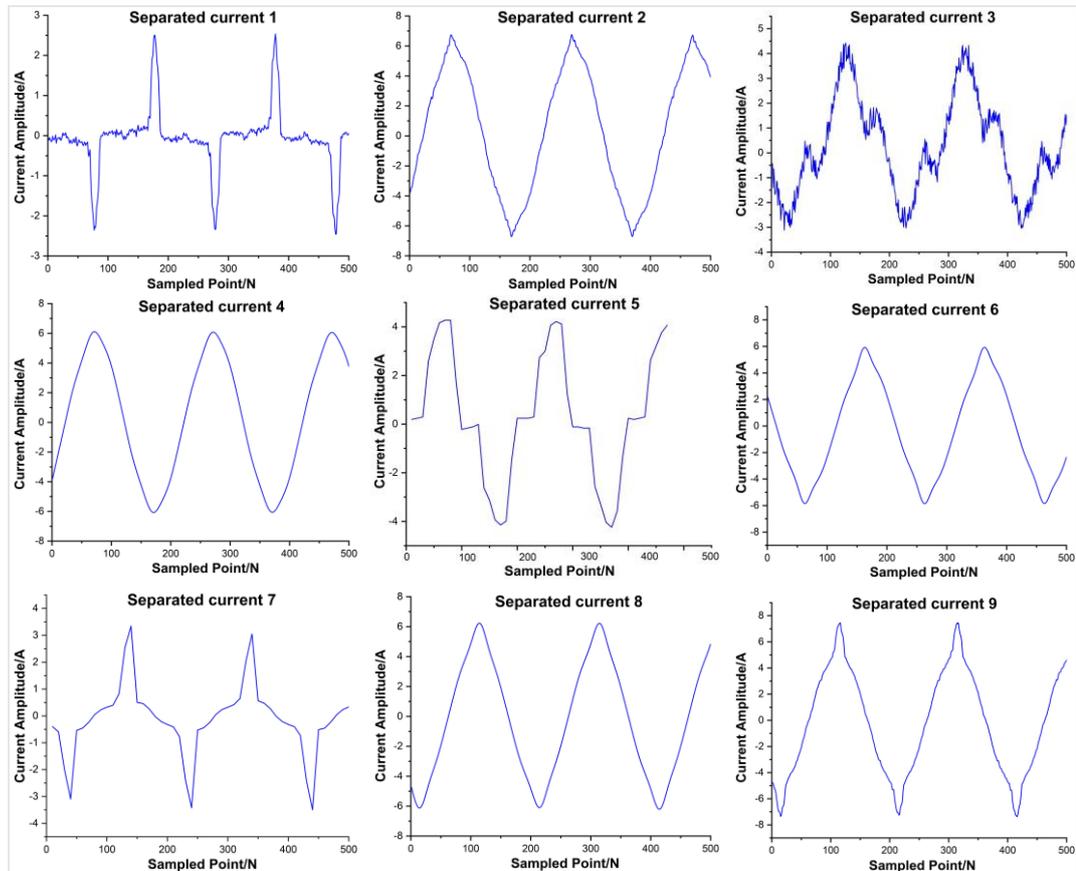
$$w(x_i,\psi_m) = \frac{1}{D(x_i,\psi_m)} \quad (5.2)$$

After normalization, the weightings of the training samples are obtained. The larger the weighting, the higher the probability that the unknown sample and training sample are from the same category of appliances, and the majority voting becomes more reliable. The basic identification through KNN is hence achieved.

### ***5.3.2 Identification Testing Based on KNN***

The above method is an improvement of the original KNN method in reference [97], with the construction of the KNN model and incorporation of the rule of majority voting. It saves storage space for the training samples set and is better suited for identification scenarios that have few appliances in each category. This thesis uses the proposed method and the method in reference [97] to identify the separated currents from House A of the UK-DALE data set. The date was 2012/11/09, and the time period was 06:00 to 22:30. A total of 59 switching events were detected and identified for this period. The 9 separated currents are displayed in Fig.5.3, and the corresponding times

to separate each sequence are given in Table 5.4. These nine sequences illustrate the identification process of the proposed KNN method.



**Fig.5. 3 Separated current working wave**

**Table 5. 4 The switching time point of 9 unknow appliance**

<b>Current Sequence</b>	<b>Separating Time</b>
Separated current 1	19:33pm
Separated current 2	19:05pm
Separated current 3	18:47pm
Separated current 4	13:23pm
Separated current 5	11:25am
Separated current 6	10:11 am
Separated current 7	07:41am
Separated current 8	07:04am
Separated current 9	20:03pm

In this identification process, the working current of the switched appliance was first separated, it is an unknown sample. The first to fourth harmonics of that separated current was then used to calculate the feature differences with training samples for the KNN model. The corresponding harmonic values are shown in Table 5.5.

**Table 5. 5 The frequency feature aggregation of separated currents**

	<b>1st</b>	<b>2nd</b>	<b>3rd</b>	<b>4th</b>
<b>Separated current 1</b>	0.307	0.017	0.224	0.010
<b>Separated current 2</b>	1.447	0.029	0.424	0.008
<b>Separated current 3</b>	0.632	0.023	0.524	0.020
<b>Separated current 4</b>	5.655	0.023	0.478	0.022
<b>Separated current 5</b>	0.611	0.026	0.504	0.018
<b>Separated current 6</b>	5.978	0.036	0.158	0.011
<b>Separated current 7</b>	0.279	0.018	0.195	0.011
<b>Separated current 8</b>	0.632	0.025	0.520	0.016
<b>Separated current 9</b>	1.543	0.042	0.493	0.017

Secondly, the Euclidean distances between the unknown sample and all the training samples were calculated using equation (2.6). The calculated values are listed in Table 5.6.

**Table 5. 6 Euclidean distance between the training sample and separated current**

	<b>TV</b>	<b>RE</b>	<b>MO</b>	<b>AC</b>	<b>LAP</b>	<b>GE</b>	<b>EK</b>	<b>VC</b>
<b>Separated current 1</b>	0.03	0.58	7.34	4.22	1.67	4.83	6.24	6.24
<b>Separated current 2</b>	0.59	0.24	6.20	3.10	1.09	3.69	5.10	5.10
<b>Separated current 3</b>	4.28	2.36	1.05	3.90	5.13	4.51	5.91	3.78
<b>Separated current 4</b>	6.31	5.36	4.85	5.35	5.88	3.31	3.23	4.78
<b>Separated current 5</b>	1.58	3.92	7.03	5.93	0.26	4.53	5.93	5.93
<b>Separated current 6</b>	5.63	4.77	3.76	1.63	1.76	1.57	1.46	2.84
<b>Separated current 7</b>	1.26	2.92	1.27	4.25	2.48	4.86	6.26	6.26
<b>Separated current 8</b>	3.28	5.57	7.01	3.90	2.13	0.42	0.34	5.91
<b>Separated current 9</b>	2.19	4.33	3.07	3.60	5.08	6.60	5.01	1.02

Thirdly, K was set as three, which means three training samples with similar frequency features as the separated current were found. Their three Euclidean distances were normalised using equation 5.2, and the results are as shown in Table 5.7.

**Table 5. 7 The normalized Euclidean distances the training sample and separated current**

	<b>TV</b>	<b>RE</b>	<b>MO</b>	<b>AC</b>	<b>LAP</b>	<b>GE</b>	<b>EK</b>	<b>VC</b>
<b>Separated current 1</b>	0.944	0.040	0	0	0.014	0	0	0
<b>Separated current 2</b>	0.156	0.712	0	0	0.130	0	0	0
<b>Separated current 3</b>	0	0.227	0.630	0	0	0	0	0.141
<b>Separated current 4</b>	0	0	0	0	0	0.368	0.377	0.255
<b>Separated current 5</b>	0.133	0.053	0	0	0.812	0	0	0
<b>Separated current 6</b>	0	0	0	0	0.301	0.338	0.363	
<b>Separated current 7</b>	0.413	0.178	0	0	0.409	0	0	0
<b>Separated current 8</b>	0	0	0	0	0.082	0.407	0.510	0
<b>Separated current 9</b>	0.259	0	0.184	0	0	0	0	0.556

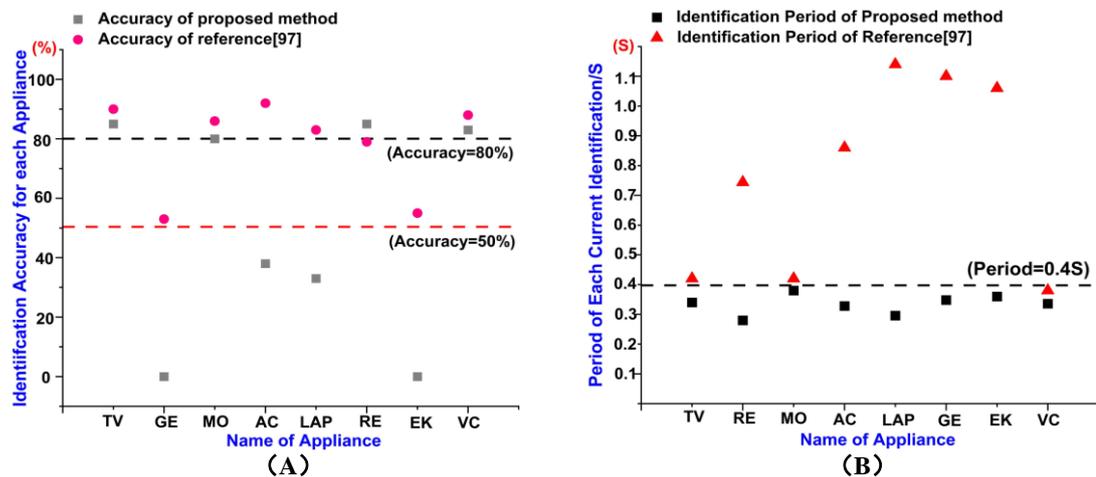
Based on the normalization results, the appliance type was then identified using the majority voting process. The identification result is listed in Table 5.8.

**Table 5. 8 Identification result based on KNN method**

<b>Separated current 1</b>	<b>TV</b>	-	-	-	-	-	-	-
<b>Separated current 2</b>	-	<b>RE</b>	-	-	-	-	-	-
<b>Separated current 3</b>	-	-	<b>MO</b>	-	-	-	-	-
<b>Separated current 4</b>	-	-	-	-	-	<b>GE</b>	<b>EK</b>	-
<b>Separated current 5</b>	-	-	-	-	<b>LAP</b>	-	-	-
<b>Separated current 6</b>	-	-	-	-	<b>LAP</b>	<b>GE</b>	-	-
<b>Separated current 7</b>	<b>TV</b>	-	-	-	<b>LAP</b>	-	-	-
<b>Separated current 8</b>	-	-	-	-	-	<b>GE</b>	<b>EK</b>	-
<b>Separated current 9</b>	-	-	-	-	-	-	-	<b>VC</b>

Unknown appliance No. 1 has been identified as the TV, No. 2 as the refrigerator, No. 3 the microwave oven, No. 5 the laptop, and No. 9 the vacuum cleaner. As for the

other unknown appliances, their features were found to be too similar to multiple appliances, resulting in similar weighting for corresponding training samples. Therefore, their identification results would require further modification and correction. Fig.5.4 compares the identification accuracy and identification period of eight appliances between the proposed method of this thesis and the method used in reference [97].



**Fig.5. 4 The comparing of identification accuracy and period between two methods**

In Fig.5.4 A, the proposed method used to identify laptop (LAP) and air-conditioning (AC) has an accuracy rate lower than 50%. This is lower than the accuracy rate achieved by the method in reference [97]. Furthermore, this proposed method could not identify the appliances with similar electric features, such as the kettle (EK) and geyser (GE). The main reason of the low identification accuracy is that when a separated current exhibits electric features of two appliances, there are no subdivided features to help differentiate them. In contrast, the method in reference [97] built a stronger granular training sample set, which required large storage space to overcome feature overlapping, but contributed to a higher total accuracy rate is higher than this proposed method. However, the identification accuracy of kettle (EK) and geyser (GE) remains around 50% for the method used in reference [97], which is still unsatisfactory for monitoring purposes.

Fig.5.4 B presents the time period used by the above two methods to complete each appliance identification. The method in reference [97] required a longer period to

identify each appliance than the proposed method for this thesis. The time required for such identification varied, depending on the type of appliance involved. For example, the identification of a laptop took 1 second to complete, but the identification of a TV took 0.4 seconds to complete. In contrast, the proposed method utilized fewer elements in its training feature set, which resulted in less time spent comparing and measuring the feature differences. Consequently, the proposed method took roughly the same amount of time to identify each appliance.

In other words, the proposed method uses a shorter identification period and required less storage space. To achieve higher accuracy for this method, the identification process for appliances with similar electric features must be modified and corrected.

#### **5.4 Modification of Identified Result Based on Appliance Operation Pattern**

During the last identification test using the KNN method, certain appliances could not be identified because their electric features were too similar, leading to interference in the identification results. As discussed in Section 5.2, for appliances with similar electric features, their non-electric features, such as the time when the appliance is switched on and its running duration, are generally distinct. While electric kettles and geysers have almost the same current features, their switching times and running durations are completely different. The appliance operation pattern can hence be applied to improve the identification accuracy. This section aims to propose the following modified method to help identify the appliances using their operation patterns.

First, the non-electric data about known appliance switching times are collected to train BPNN. After completing the training, the operation pattern curves of known appliances can then be obtained.

Next, three additional weeks of appliance switching data are gathered, which includes the times of the appliance being switched on and each appliance running period, to validate the appliance operation pattern curve obtained.

Lastly, the running and switching conditions of the unknown appliance are compared with the operation pattern curves of the known appliances. If the unknown

appliance operations did not match with the operation pattern curves of a given known appliance, the unknown appliance is not the same type as the known appliance. Conversely, the unknown appliance type can be identified if its operation patterns are consistent with that of a known appliance.

#### 5.4.1 BPNN Principle for Operation Pattern Regression

As outlined in Section 5.2, a longer statistical period of non-electric data is vital to obtain accurate appliance operation patterns and switching probability distribution. However, this can be difficult to realize in practical applications. Therefore, BPNN is used to process the short-term statistical data to obtain accurate appliance operation patterns and switching probability distribution curves.

BPNN is a feed-forward network model with a multilayer system [100] [69], which comprises an input layer, a hidden layer, and an output layer. The input layer receives data that holds significance in the real world. The hidden layer processes the input data based on the corresponding activation function. The output layer exports the network's response to the input data. Each layer contains multiple neurons which are connected between layers via connection weights. The topological diagram of a typical BPNN structure is shown in Fig.5.5.

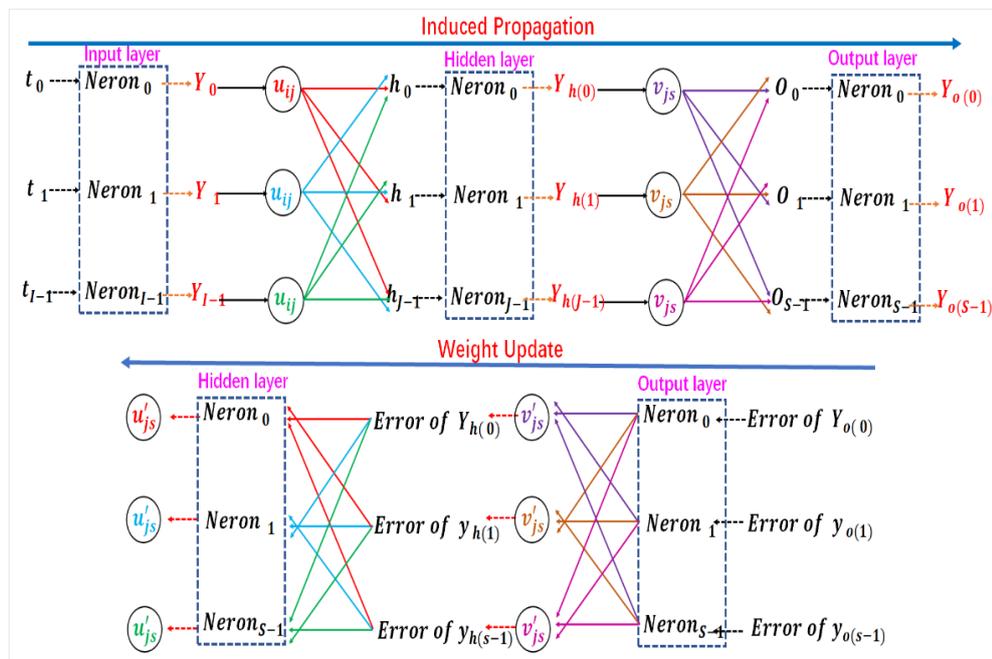


Fig.5. 5 BPNN structure

Fig.5.5 shows how the BPNN performs the induced propagation and the weight update.  $I, J$  and  $S$  are the number of neurons in the input layer, the hidden layer, and the output layer, respectively. Each layer has its independent input and output.  $u_{ij}$  denotes the connection weighting between  $i$ th neuron of the input layer and  $j$ th neuron of the hidden layer.  $v_{js}$  is the connection weighting between  $j$ th neuron of the hidden layer and  $s$ th neuron of the output layer respectively.

The quantity of input layer neurons is determined by the nature of the BPNN's input data. An activation function is not used in the input layer, so the input  $t_i$  of neurons in the input layer is mapped to the hidden layer directly.

As for data processing in the hidden layer, the number of neurons in the hidden layer is set manually and can adjust according to different application requirements. Every neuron in the hidden layer has independent input data which is derived from each input layer neuron. The input of the hidden layer's neurons is  $h_j$ . The output of the hidden layer is  $Y_{h(j)}$ . Then  $Y_{h(j)}$  is mapped to the output layer.

As for data processing in the output layer, the number of neurons is set manually according to requirements. Based on the output data from the hidden layer, the output layer feeds back its response to the BPNN's input data. The input of the output layer's neuron is  $O_s$ . The output of the output layer's neuron is  $Y_{O(s)}$ . The induced propagation is completed when obtaining the output of each neuron in the output layer.

The network error in the BPNN is then obtained by calculating the difference between the expected input data and the induced response. If the error is large, the network error will be decreased through connection weight updates.

The BPNN is constructed once the activation function and the number of neurons in each layer have been fixed. The induced propagation and connection weighting updates are repeated until the network error falls within an acceptable pre-determined range. This is the training of the BPNN. If a trained BPNN was obtained, the BPNN will be used in the corresponding application.

#### ***5.4.2 Operation Pattern Regression Based on BPNN***

This section presents the building of BPNN and the learning strategy for operation

pattern regression. First, the dataset being used for training the BPNN is constructed. Secondly, depending on the size of training dataset, the network layers and the number of neurons in each layer are determined. Thirdly, the BPNN is trained using a consecutive update of connection weight. Finally, the appliance operation pattern curve can be regressed according to the designed learning strategy.

#### 5.4.2.1 Training data set building

The size of training data affects the convergence of training. The potential dataset on appliance operations is huge, which can include data such as running period, startup time, and switching times. The more data is being used, the more accurate the operation patterns obtained, but this may affect the convergence of training [101] [102]. To improve convergence speed while maintaining data diversity, the one day period was divided into 24 time periods. The number of times a known appliance was switched on in each period was counted respectively, using the UK-DALE data set. By considering the number of appliance switching times, the appliance running periods and switching time points are captured within one single dataset. Four appliances, namely the geyser, TV, laptop, and microwave oven (MO), were selected to train the BPNN as shown in Table 5.9:

**Table 5.9 The number of switching times of four appliance [UK]**

<b>Period</b>	<b>AM :1-2</b>	<b>AM :2-3</b>	<b>AM :3-4</b>	<b>AM :4-5</b>	<b>AM :5-6</b>	<b>AM :6-7</b>	<b>AM :7-8</b>	<b>AM :8-9</b>	<b>AM: 9-10</b>	<b>AM: 10-11</b>	<b>AM: 11-12</b>	<b>AM:12- -PM:1</b>
<b>Geyser</b>	6	7	8	9	9	8	8	8	8	10	9	9
<b>TV</b>	0	0	0	0	0	0	0	0	0	0	1	3
<b>Laptop</b>	0	0	0	0	0	0	0	3	7	7	5	2
<b>MO</b>	0	0	0	0	0	0	0	0	1	0	0	1
<b>Period</b>	<b>PM :1-2</b>	<b>PM :2-3</b>	<b>PM :3-4</b>	<b>PM :4-5</b>	<b>PM :5-6</b>	<b>PM :6-7</b>	<b>PM :7-8</b>	<b>PM :8-9</b>	<b>PM: 9-10</b>	<b>PM: 10-11</b>	<b>PM: 11-12</b>	<b>PM:12- AM:1</b>
<b>Geyser</b>	10	10	10	9	7	10	9	12	11	13	12	9
<b>TV</b>	0	0	1	1	2	5	4	4	3	3	1	0
<b>Laptop</b>	1	1	2	1	0	4	7	6	8	6	2	1
<b>MO</b>	0	0	0	0	0	0	1	1	3	0	0	0

In Table 5.9, the first six rows show the 24 time periods of the day: 1:00 am to 2:00 am is the 1st period, 2:00 am to 3:00 am is the 2nd period, and 12:00 pm to 1:00

am is the 24th period. The remaining rows show the number of times each device was switched on during each corresponding period.

The data in Table 5.9 was subsequently converted into a matrix for training purposes. The first row of the matrix displays the 24-hour time periods, and the second row displays the number of times an appliance was switched on in each corresponding period. The data matrixes 1 to 4 are then used to train the network on the operations of the geyser, TV, laptop, and microwave oven.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 4 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 6 & 9 & 2 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}$$

Matrix 1: Training data set of Geyser

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 4 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 6 & 9 & 2 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}$$

Matrix 2: Training data set of TV

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 4 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 6 & 9 & 2 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}$$

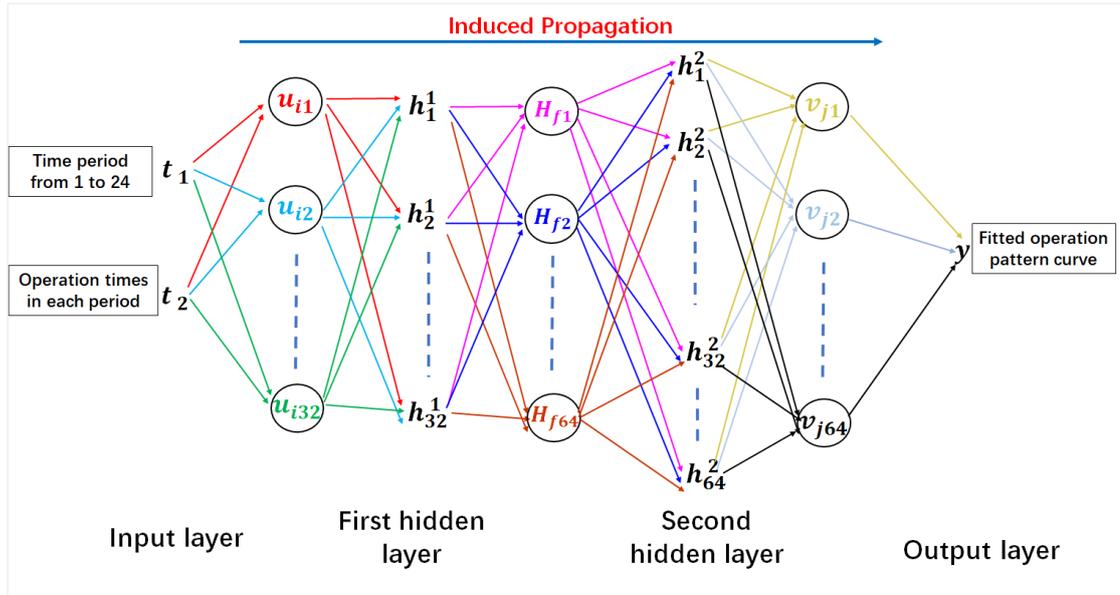
Matrix 3: Training data set of Laptop

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 4 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 6 & 9 & 2 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}$$

Matrix 4: Training data set of Microwave Oven

#### 5.4.2.2 BPNN Training

After obtaining the training data, the BPNN for operation pattern regressing is constructed and trained next. Depending on the size of the training dataset used, the BPNN is constructed accordingly, in this case, it consisted of one input layer, two hidden layers and one output layer. The input layer had 2 neurons. The first hidden layer had 32 neurons, and the second hidden layer had 64 neurons. The output layer had 1 neuron. The proposed BPNN structure and the general data process is illustrated in Fig.5.6.



**Fig.5. 6 BPNN structure for operation pattern training**

Once the layer and neuron structures of the BPNN were established, the training data was input into the network, and a response from the network was obtained. Using the geyser as an example, the process of obtaining a response was as follows:

First, the geyser was switched on six times in the first period of the day, from 1 am to 2 am, so  $\begin{bmatrix} 1 \\ 6 \end{bmatrix}$  was input into the network. and the output of the BPNN is obtained using equations (5.3) to (5.7).

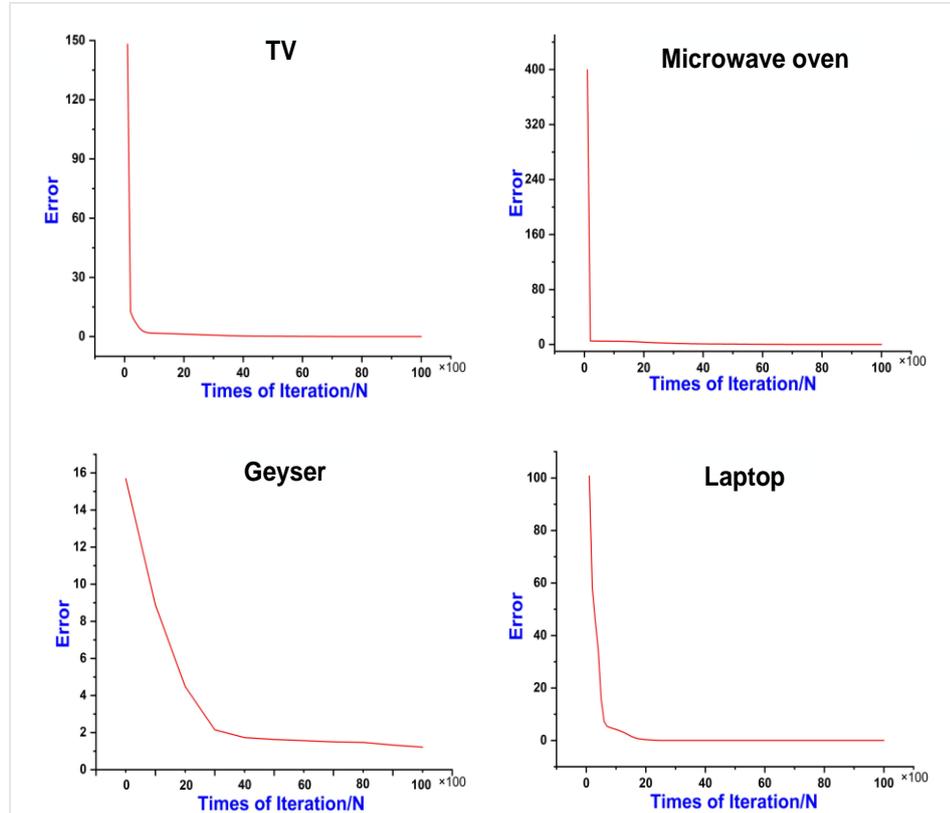
Next, the geyser was switched on six times in the second period of the day from 1am to 2 am, so  $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$  was input into the BP network. The corresponding output of the BPNN was then obtained.

This process was repeated for all 24 columns of training data which were inputted into the network one after another to obtain the 24 outputs of the BPNN.

Using these 24 outputs of the BPNN, the training data was regressed to an operation pattern curve.

Lastly, the error between the regressed operation pattern curve and the real data is calculated using equations (5.8) to (5.11), then the connection weighting is updated to decrease the error. The connection weighting was updated 10000 times. Fig.5.7 shows the convergence trend of network error for the four BPNNs during the training

process. As the number of updates increased, the error decreased to almost zero. Furthermore, the network error of all four BPNNs decayed exponentially, which suggests that the activation function and structure of the network are both appropriate.



**Fig.5. 7 BNPP error trend during training process**

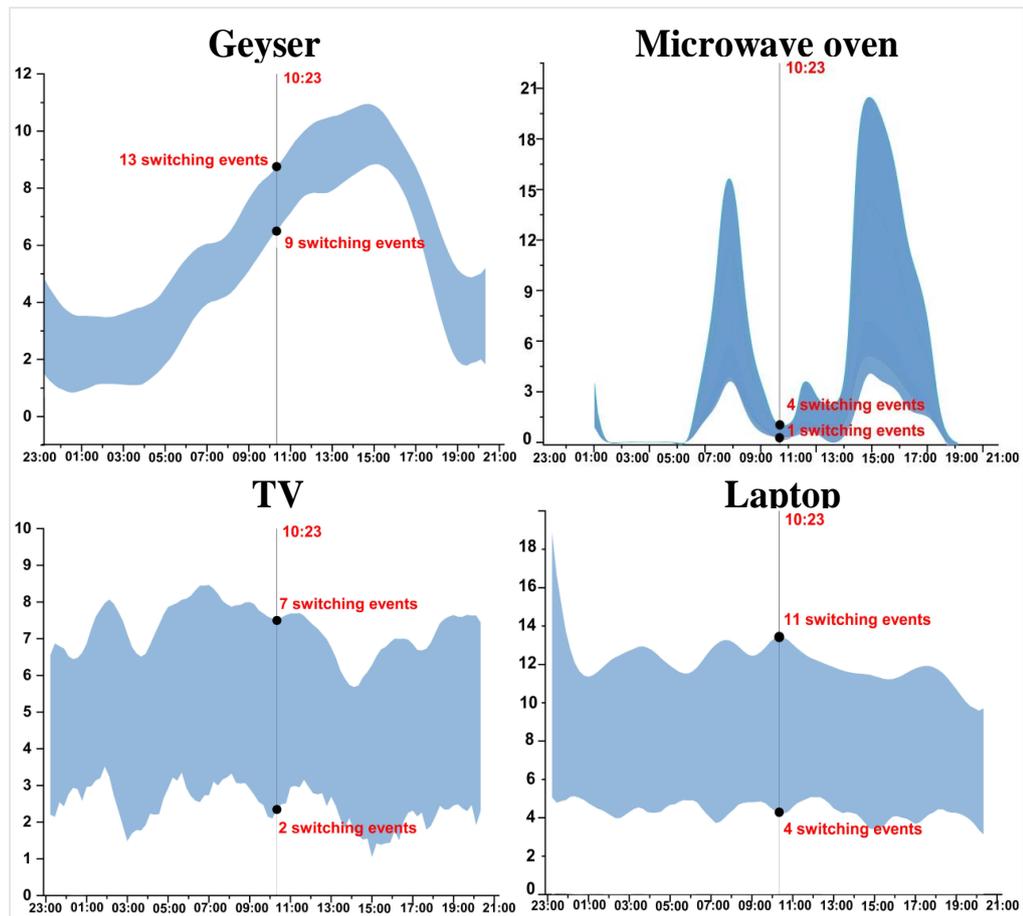
When the network error was close to 0, the training of the network is completed. A mature BPNN is constructed.

#### 5.4.2.3 The Appliance Operation Pattern Regression

Once the training of BPNN is completed, the operation pattern curve regressed by the BPNN will closely approximate to the actual running situation. However, this operation pattern curve only conforms to the running situation of the specific day on which the training data was extracted. Hence, the obtained operation pattern curve cannot reflect a universal appliance running situation. Take for instance the UK-DALE dataset which includes 786 days of switching data. Based on the regressed operation

pattern curve, it is difficult to accurately evaluate and reflect the possibility of appliance switching patterns across the entire 786 days. Therefore, it is necessary to fit and regress multiple days' operation pattern curves, subsequently analysing and evaluating the possibility of appliance switching under multiple operation pattern curves.

A random extraction of 40 days' data was taken from the 786 days, focusing on the switching data of geyser, TV, laptop and microwave. Each day's switching data was regressed to obtain an operation pattern curve which is then placed under the same coordinate system. The ordinate represents the number of times of appliances switching, the abscissa represents the time of the day. Fig.5.8 shows 40 regressed operation pattern curves obtained.



**Fig.5. 8 The operation pattern curves sets**

In Fig.5.8, the abscissa is the time of day, the ordinate is the number of switching events, the 40 regressed appliance operation pattern curves form a specific zone, the upper edge of the area represents the maximum number of switching events at a given time, the lower edge of the area represents the minimum number of switching events at a given time. Take for example the switching of geyser over the 40 days. Of all the geyser switching events, a minimum of 9 and a maximum of 13 of them occurred at 10.23 am. It is highly probable that the switching events that occurred at any time of any day fall within this zone. Therefore, based on the obtained zone and operation pattern curves, the average values of the 40 curves can be calculated, and a new operation pattern curve is thus derived. Using this new operation pattern curve, the total number of switching events within one day can now be calculated. The number of switching events at each minute is divided by the total number of switching events to calculate the switching probability of the appliance at every minute of the day. The appliance operation pattern curves with the switching probability are shown in Fig.5.9 to Fig.5.12, which gives the probability of the appliance switched at each minute of the day.

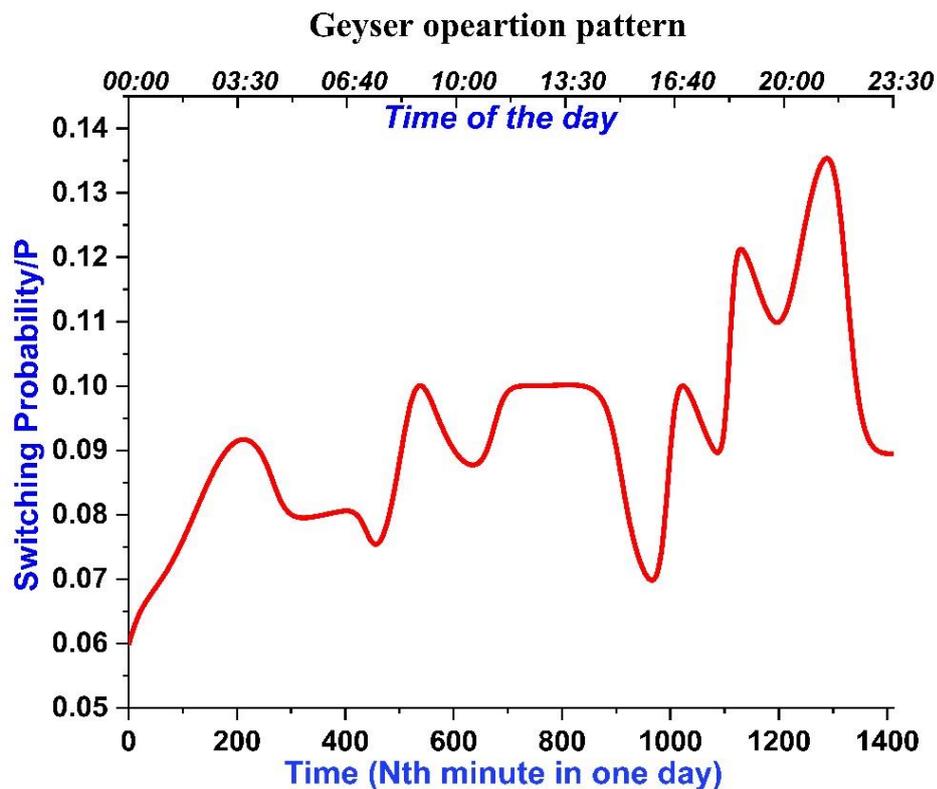


Fig.5. 9 The general operation pattern curve of a geyser

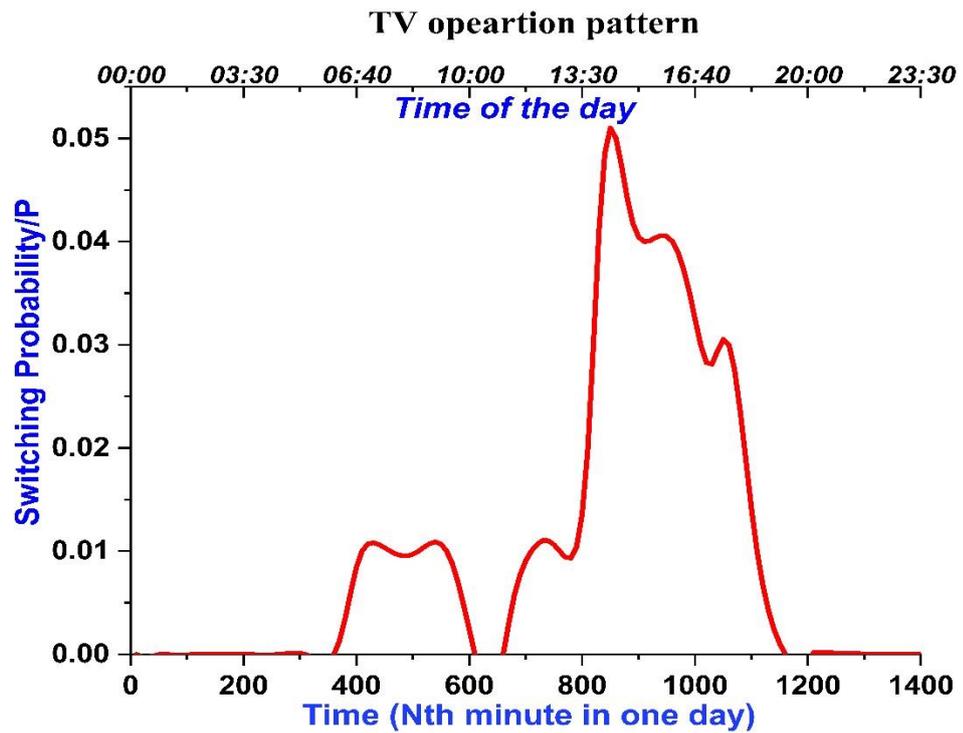


Fig.5. 10 The general operation pattern curve of a TV

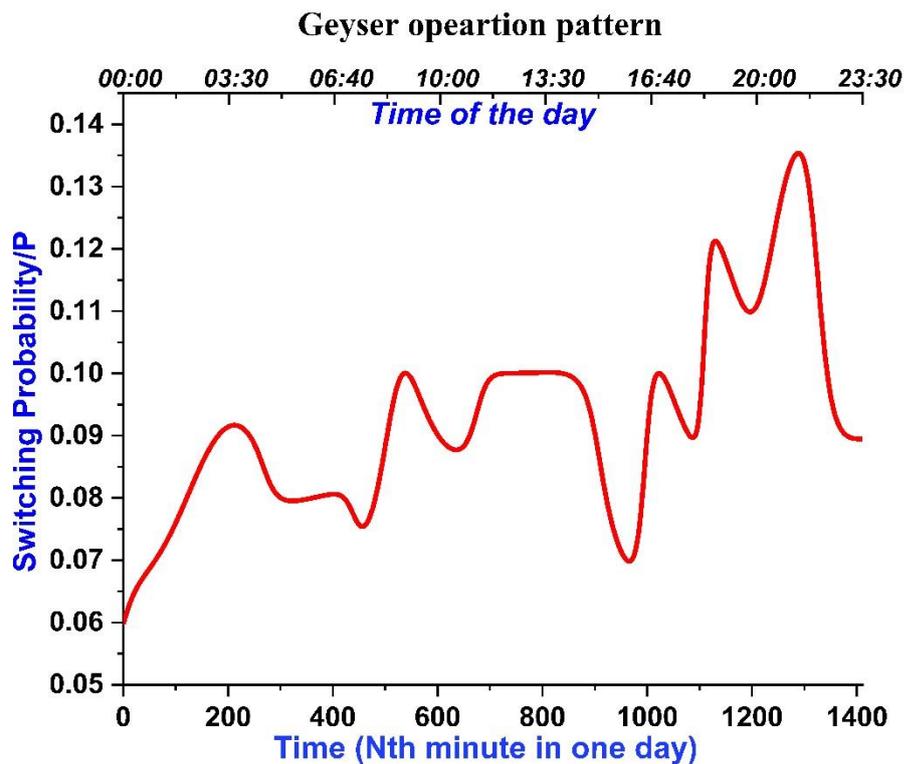


Fig.5. 11 The general operation pattern curve of a laptop

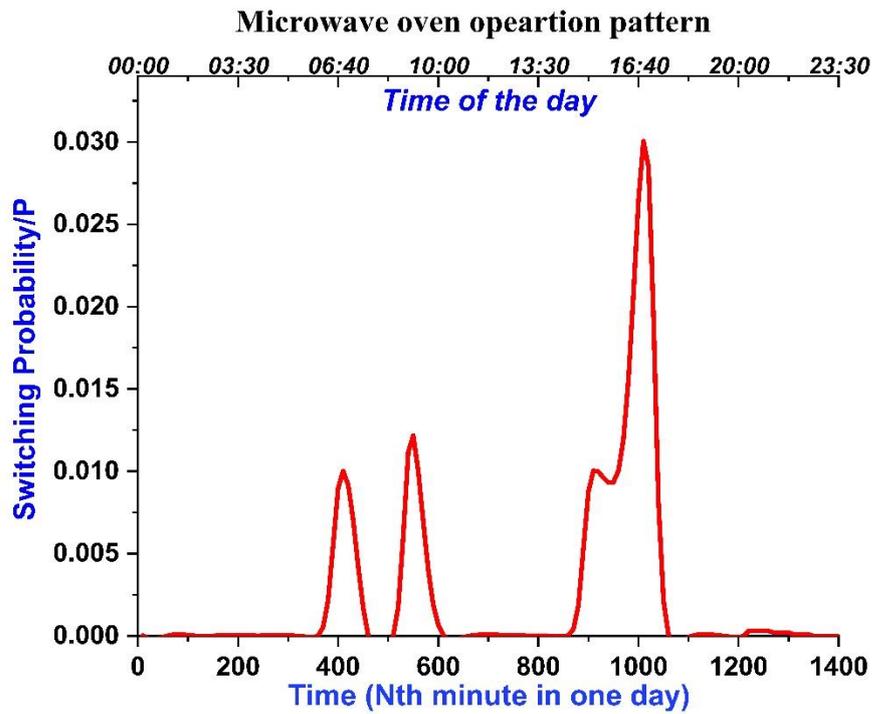


Fig.5. 12 The general operation pattern curve of a microwave oven

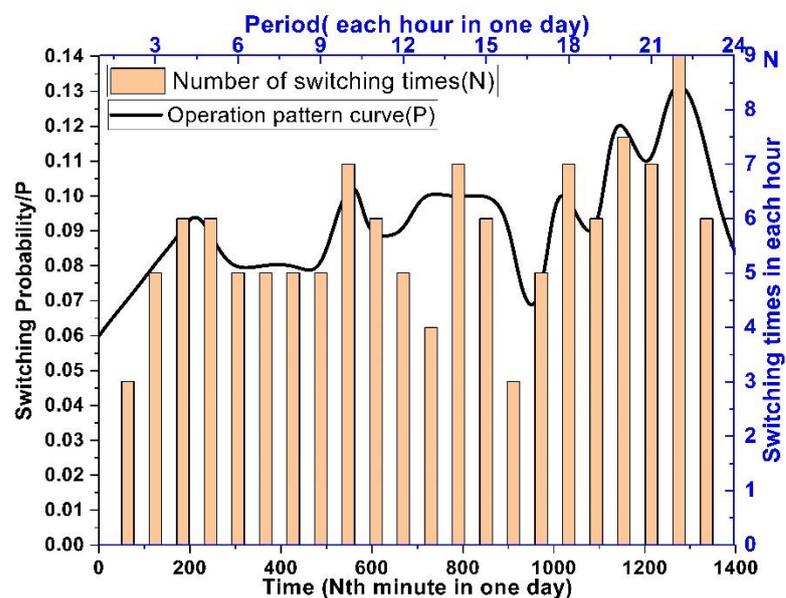
### 5.4.3 Verification and Testing of Operation Pattern Curve

When the operation pattern curve and appliance switching probability were obtained, the others three weeks' data was used to prove and test the operation pattern curve and switching probability. The number of switching events of a geyser, TV, laptop, and microwave oven (MO) in each hour was counted, it is also measured from the UK-DEAL data sets, Table 5.10 shows the average number of switching times of the four appliances during each hour.

Table 5. 10 The number of switching times of four appliances [UK]

Period	AM :1-2	AM :2-3	AM :3-4	AM :4-5	AM :5-6	AM :6-7	AM :7-8	AM :8-9	AM :9-10	AM: 10-11	AM: 11-12	AM:12-PM:1
Geyser	5	7	9	8	9	7	9	7	8	11	8	8
TV	0	0	0	0	0	0	0	0	0	0	1	3
Laptop	0	0	0	0	0	0	0	3	7	7	5	2
MO	0	0	0	0	0	0	0	0	1	0	0	1
Period	PM: 1-2	PM: 2-3	PM: 3-4	PM: 4-5	PM: 5-6	PM: 6-7	PM: 7-8	PM: 8-9	PM: 9-10	PM: 10-11	PM: 11-12	PM:12-AM:1
Geyser	10	10	9	9	8	11	7	10	11	12	12	11
TV	0	0	1	1	2	5	4	4	3	3	1	0
Laptop	1	1	2	1	0	4	7	6	8	6	2	1
MO	0	0	0	0	0	0	1	1	3	0	0	0

Fig.5.13 compares the geyser's operation pattern curve and the actual statistical switching times of the geyser in each hour. In Fig.5.13, the line chart and the histogram chart are compared together. As for the histogram chart, the right ordinate is the number of times the appliance is switching, the top abscissa is each hour of the day, under the corresponding axis, and the histogram is the actual statistical times of the geyser running in each hour. As for the line chart, the left ordinate is the probability being the appliance has been switched on, the bottom abscissa is each minute of the day (1440 minutes), under the corresponding axis, and the black line shows the appliance operation pattern curve, which represents the probability of the appliance switching on.



**Fig.5. 13 Geyser regression operation pattern curve and real statistical times**

In Fig.5.23, comparing the number of switching times and the operation pattern curve that the geyser was operated, most of the geyser switching-on events happened when the probability was high in the operation pattern curve, which means the switched-on probability distribution curve is reliable, and the operation pattern of the appliance is completely regressed.

From Fig.5.14 to Fig.5.16, they compare the number of a given appliance switching times and the appliance's operation pattern curve. For all the appliances tested, the number of appliances switching times follows the curve trend of the regressed operation pattern. In the periods in which it was determined as a high

probability that appliances were switched on, the number of appliance switching times was also high. By contrast, in the period in which it was determined as less likely that an appliance was switched on, the number of times was also low. Therefore, the appliance running coincides with the trend of the operation pattern curve.

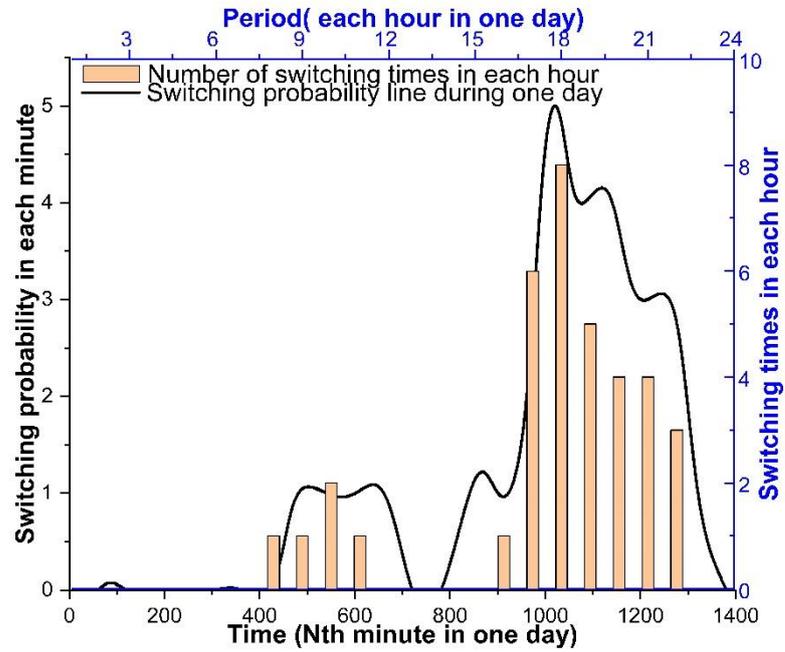


Fig.5. 14 TV regression operation pattern curve and real statistical times

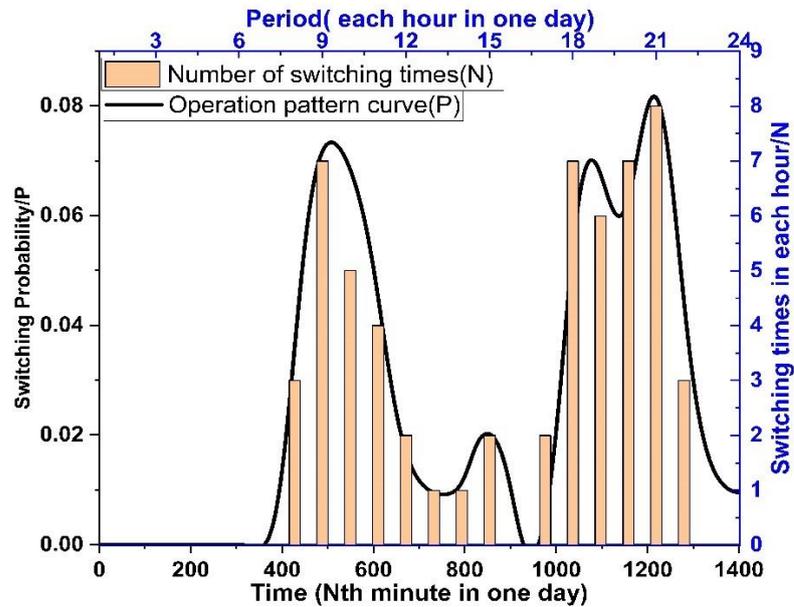
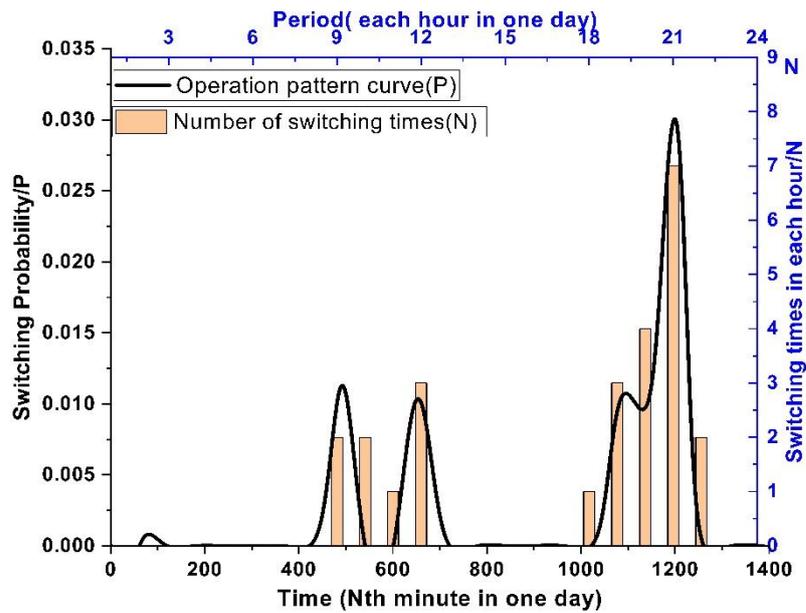


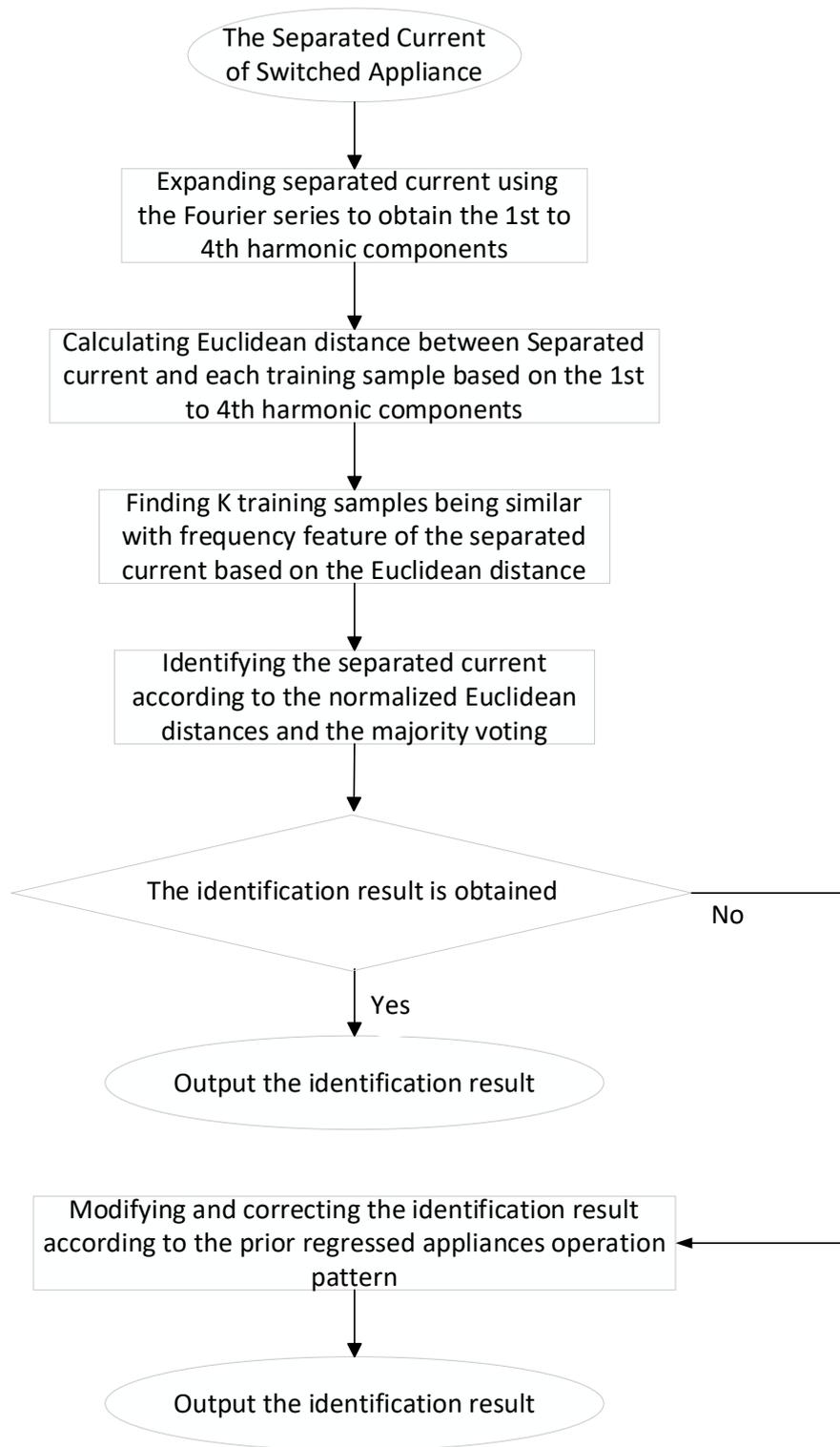
Fig.5. 15 Laptop regression operation pattern curve and real statistical times



**Fig.5. 16 Microwave oven regression operation pattern curve and real statistical times**

#### ***5.4.4 Identification Correction Testing***

The operation pattern curve gives the regularity and possible future trajectory of an appliance running. It can be used as a reference to check whether identification is correct or not. Therefore, the whole identification process includes two parts, the first part is the basic identification according to the KNN method, which can identify the appliance with unique electric features. The second part is the correction and modification of the identification results. The correction and modification are completed under the appliance operation pattern. The flow chart Fig.5.17 shows the whole identification process.



**Fig.5. 17 The flow chart of the whole identification process**

In identifying about nine separated currents in Fig.5.3, the separated currents 4, 6, 7, and 8 can be identified as two different appliances, which is obviously

problematic. However, the switched-on probabilities of different appliances are different at specific time points, derived as they are from the operation pattern of each different type of appliance. Using this data to correct the KNN method's findings allows for more secure identification.

Separated current 4 provides a example. The appliance was switched on at 13:23 pm, the 803rd minute of the day. Fig.5.18 and Fig.5.19 show the probability of the geyser and kettle being switched on at 13:23 pm:

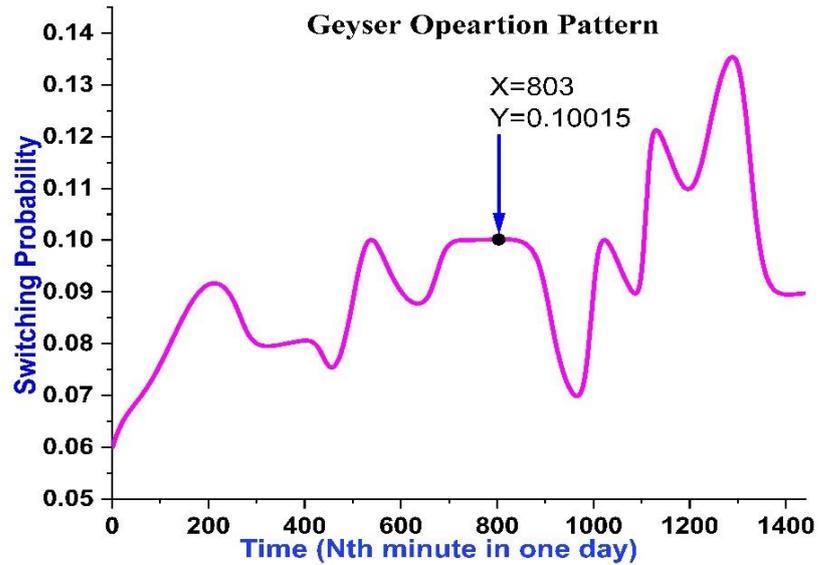


Fig.5. 18 Switched-on probability of the geyser in 803rd minute

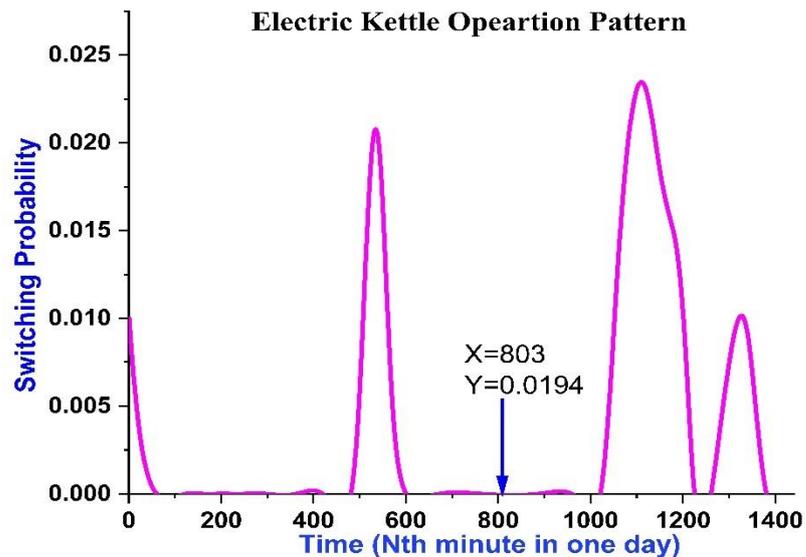


Fig.5. 19 Switched-on probability of the kettle in 803rd minute

From these figures, the geyser has a higher switched-on probability than the kettle at the corresponding point in time. This appliance is thus extremely unlikely to be the kettle. The separated current 4 belongs to the geyser. The mistake in identity attribution in the KNN method is corrected by using the operation pattern curve. As for the separated current 7 could be identified as either the TV or laptop. Fig.5.20 and Fig.5.21 shows that the laptop was likely being operated at the corresponding time:

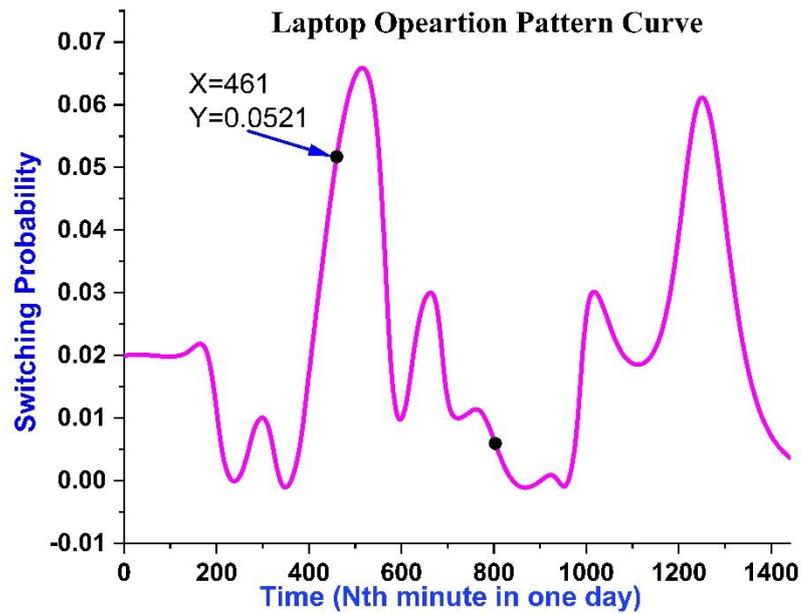


Fig.5. 20 Switched-on probability of the laptop at the 461st minute

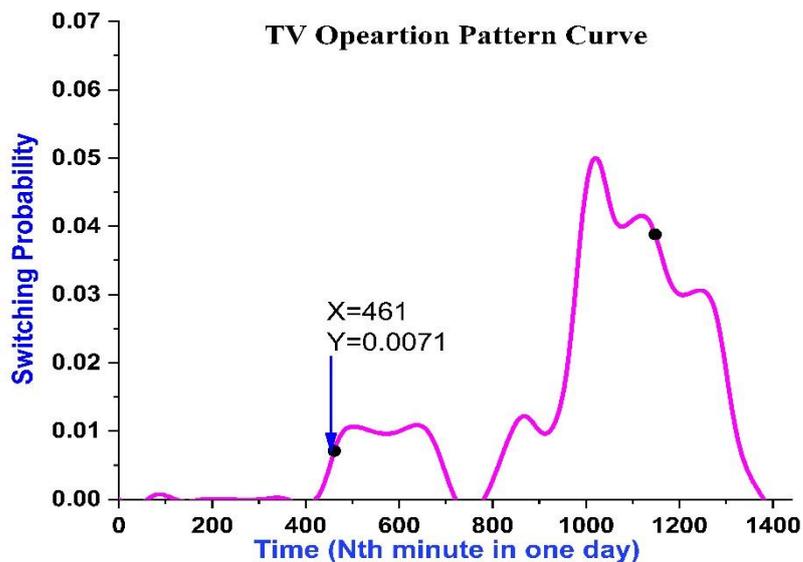


Fig.5. 21 Switched-on probability of the TV at 461st minute

Given both the appliance working currents and the appliance operation pattern curves, the identification of nine separated currents was obtained and is shown in Table 5.11.

**Table 5. 11 The nine separated current identification result**

<b>Current Sequence</b>	<b>Identified result</b>
Separated current 1	TV
Separated current 2	Refrigerator
Separated current 3	Microwave oven
Separated current 4	Geyser
Separated current 5	Laptop
Separated current 6	Electric kettle
Separated current 7	Laptop
Separated current 8	Geyser
Separated current 9	Vacuum cleaner

## 5.5 Summary

This chapter aims to identify the working current of switched household appliances. The KNN method for classifying and identifying appliances is first outlined. A general process is then presented for achieving the current identification of a switched household appliance using the KNN method. The proposed process thus completes the basic identification of appliance type.

In order to overcome the mistakes caused by overlaps in the appliances' electric features, the BPNN regressed the operation pattern curve of the known appliance, which profiles the switched-on probability distribution curve for each individual appliance. This appliance operation pattern serves as a reference to correct and modifies the identification of unknown appliances. Once the methodology and construction of BPNN for regressing appliance operation patterns are completed, the learning strategy to obtain a general operation pattern is designed. Using the BPNN and learning strategy, the general operation pattern is regressed to verify and refine the identification results of the KNN model.

Once the identification model is established, it is integrated with the switching

detection and current separation model to form a complete NILM system. This integrated system can be applied to real-world power environments, and the results can be processed. The NILM system was applied in the following three ways: (1) a three-day trial using recorded data from a real-world power environment; (2) a one-day trial application conducted in four households; (3) a 90-days application conducted in a single household. The next chapter shall present these in greater details.

## Chapter 6 Application of Proposed Method

### 6.1 Introduction

According to the proposed method, two sets of tests are performed. The first one is an offline test but uses life-recorded data. The second test is online and is a direct application to domestic homes.

In the first offline test, the physical monitoring device is constructed, which is used to capture live data, the duration of recorded data is up to three days. Then the recorded data was processed.

In the second online test, the monitor device is connected to four homes. The duration of the monitor is one day. Then the monitoring device was connected to one home to monitor the power consumption over three months (ninety days). This last-long period test can be used in the future to improve power network stability for demand-side management.

### 6.2 Monitoring Hardware

As the proposed method utilizes high-frequency data, specialized data acquisition devices need to be designed. The monitoring hardware are shown in Fig.6.1.

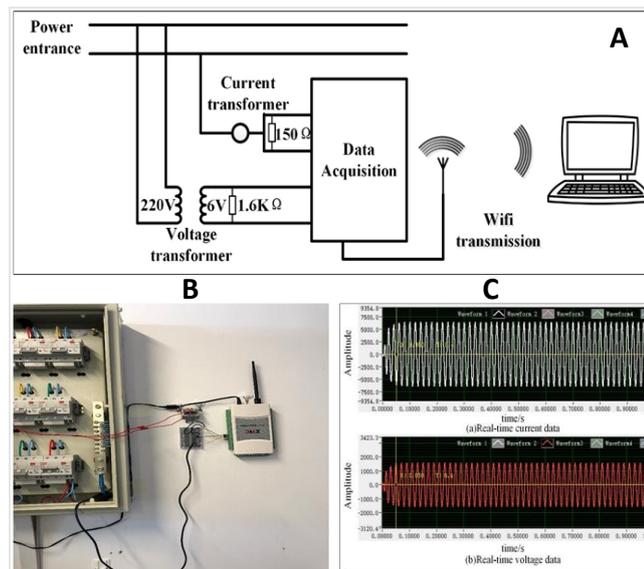


Fig.6. 1 Schematic diagram of device and system

Fig.6.1A shows the schematic diagram of such a monitoring device, while Fig.6.1B shows the prototype hardware, and Fig.6.1C displays the sampled electric data.

In Fig.6.1 A, the monitoring device comprises a data acquisition device, current transformer and voltage transformer that possess high sampling rate, high resolution and wireless transmission capabilities.

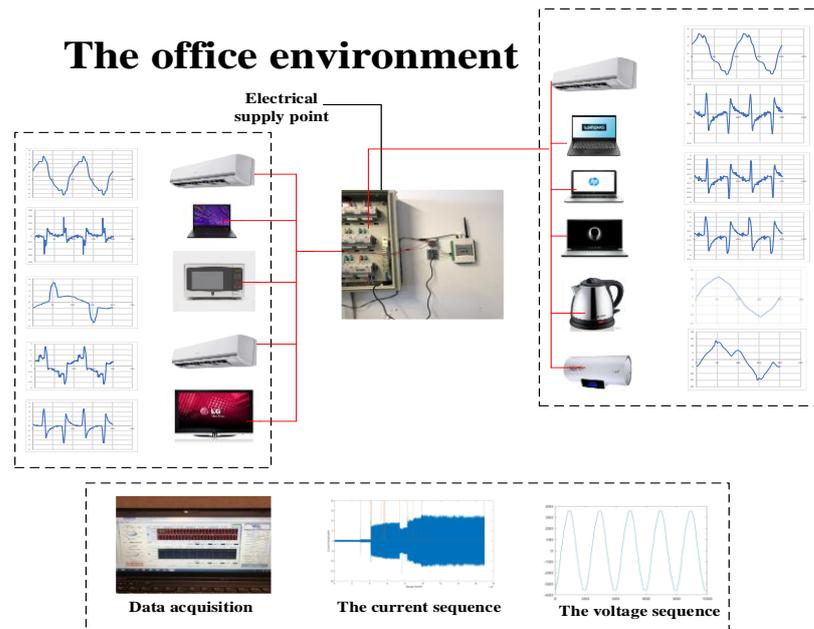
As seen in Fig.6.1B, the data acquisition device is installed at the electricity supply point. Generally, low current is employed in signal processing. However, both the current of the house and individual household appliances involve high currents. If the house current is processed and monitored directly, it would be difficult to achieve a high accuracy monitoring. Therefore, current and voltage transformers are used to act as the bridge between the power system and the monitoring system, converting the high current into lower current for easier processing of the current. Concurrently, the acquired data is transmitted to the computer through Wi-Fi for analysis.

In Fig.6.1 C, the obtained current and voltage sequence are undistorted and integrated sine waveforms.

Once the data acquisition device is constructed, it is applied in three scenarios to verify the effectiveness of individual appliance level monitoring. The first one, based on the data acquisition device, 3 days' data are collected and recorded from a student office at one university (Beijing), there are four students in this office, the collection date is from the 26/08/2020 to 28/08/2020. The second one, the data acquisition device is used in four normal domestic homes, which covers 24 hours, the application date is 07/16/2020, the homes are in a residential block at Taiyuan, the number of people in one home is from three to five; The third one, 90 days' on-line application is based on 1 domestic home for long time monitoring, the application date is from 01/06/2020 to 01/09/2020, the home is locating at Taiyuan, the number of people in home is three.

### **6.3 Off-line Application Based on 3 days' recorded Data**

Based on the student office, the monitoring device is installed at the electricity supply point, Fig.6.2 shows the office and monitoring environment, the office keeps away from noise sources, the harmonic intrusion is unserious, and the SNR of the current signal is high.



**Fig.6. 2 Monitoring environment of office**

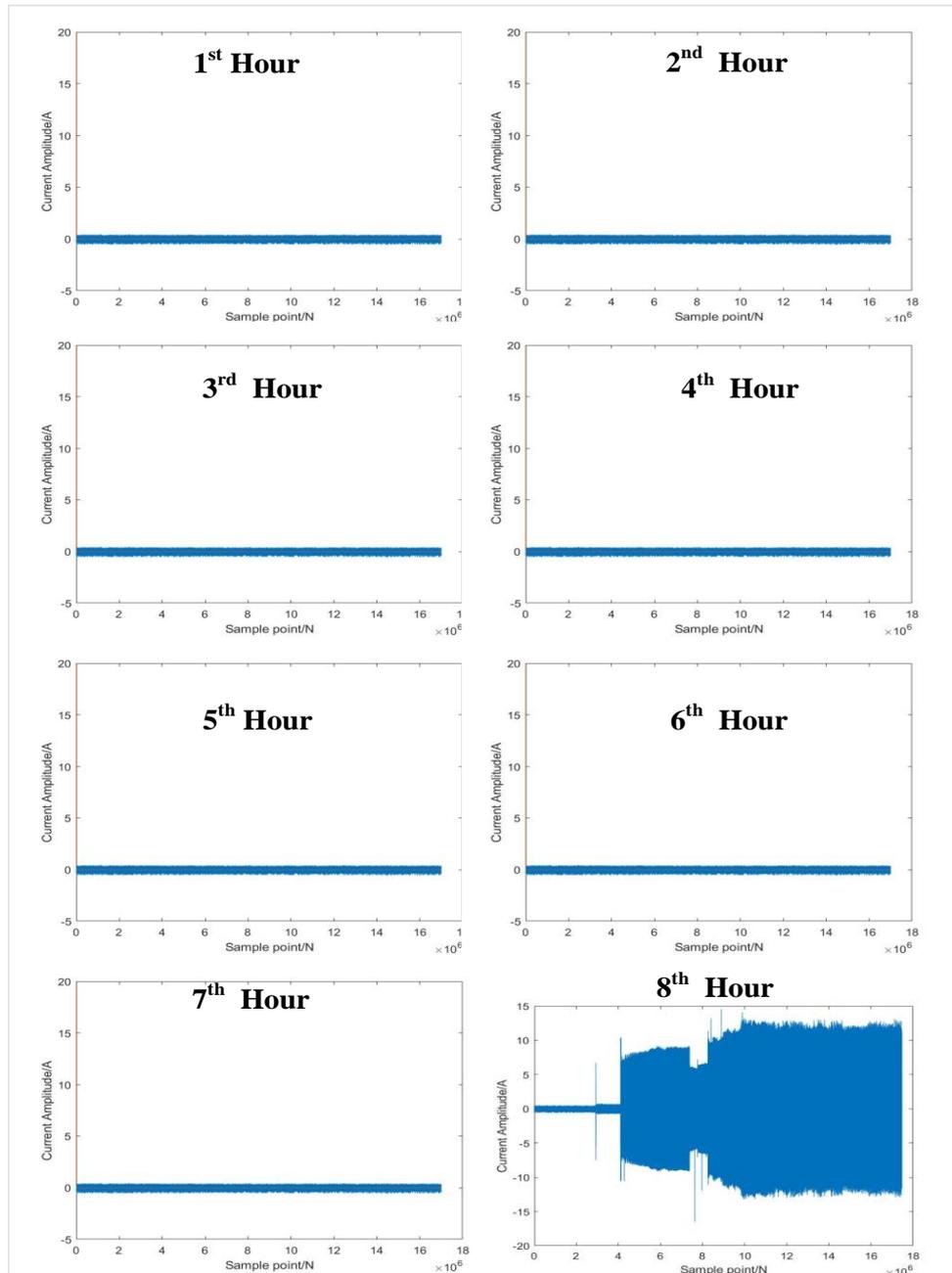
Four postgraduate students work in the office. The office includes the working zone, meeting zone, and resting zone. The number of appliances in the office is less. The original current and voltage data are completely recorded from the electricity supply entry point, the sampling frequency is 5kHz, and there is not any pre-processing of the data, such as noise reduction and data smoothing. Then the recorded data was processed by the proposed method in Chapter 4 and Chapter 5. The main monitoring objects are eleven appliances: four laptops, three air-conditioners, a kettle, a geyser, a microwave oven, and a TV. All appliance details are shown in Table 6.1.

**Table 6. 1 Appliance name and type**

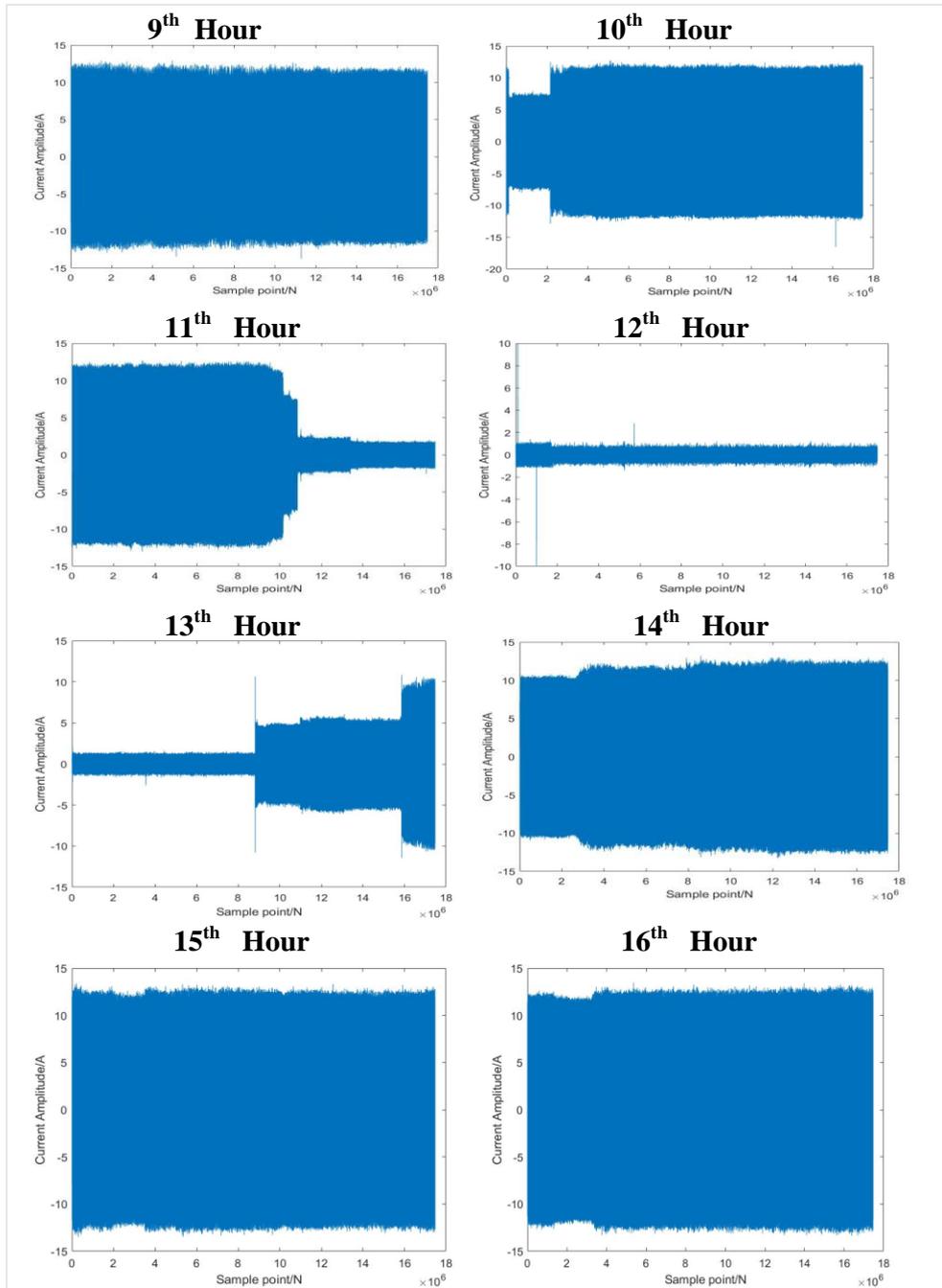
<b>Appliance Name</b>	<b>Brand</b>	<b>Abbreviation for monitoring</b>
Air-conditioning	Haier	AC-A/B/C
Laptop	Lenovo	L-A
Laptop	HP	L-B
Laptop	Dell	L-C
Laptop	Lenovo	L-D
Microwave oven	SHARP	MO
Kettle	SUPOR	K
TV	LG	TV
Geyser	Haier	G

### 6.3.1 Appliance switching detection and type identification

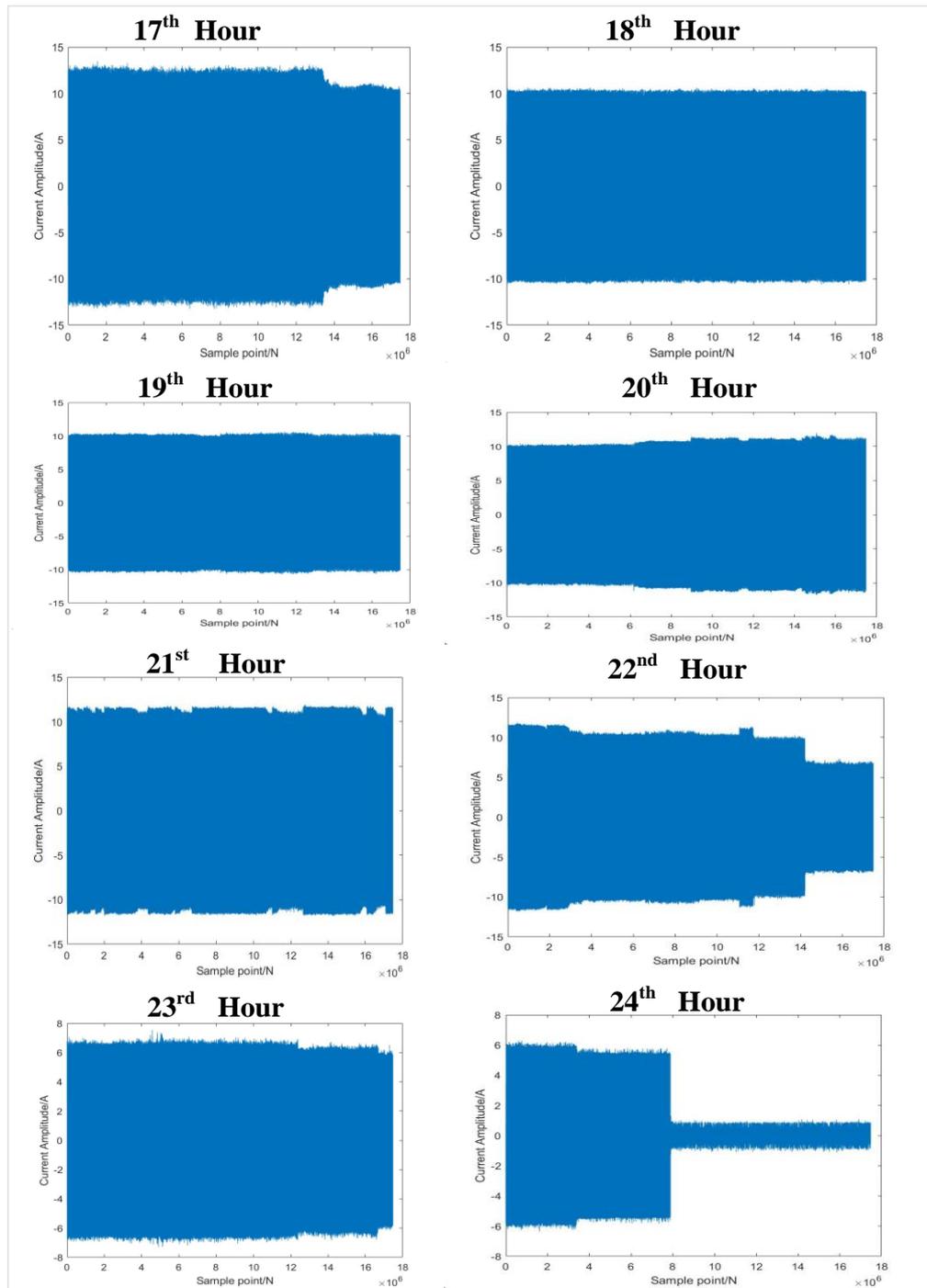
The sampling frequency is 5kHz, so around  $1.8 \times 10^7$  data were recorded every hour. Fig.6.3 to Fig.6.5 shows the first day's recorded data, another two days' data are in the appendix I.



**Fig.6. 3 Total current of office from 1st hour to 8th hour**



**Fig.6. 4 Total current of office from 9th hour to 16th hour**

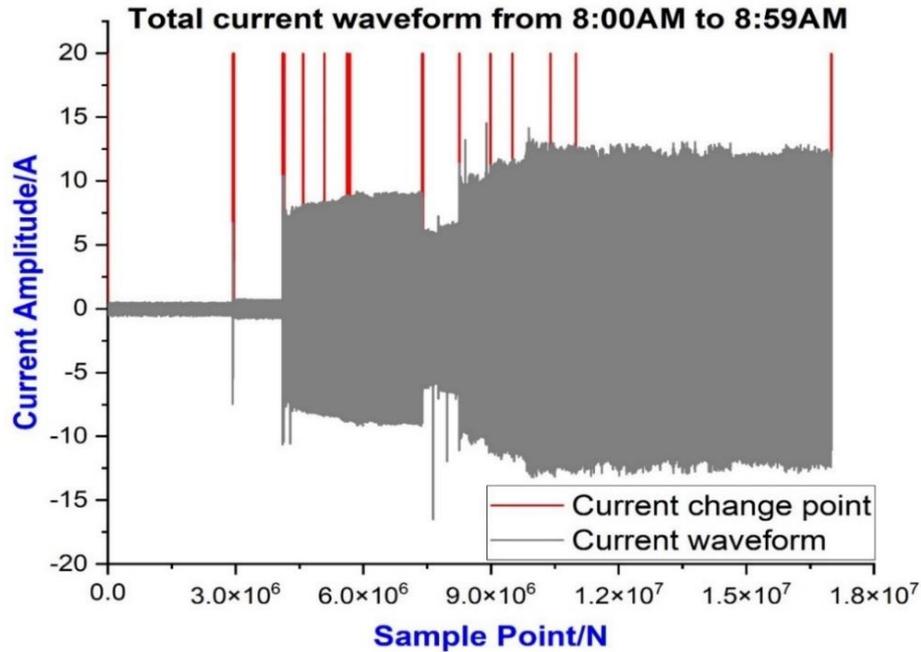


**Fig.6. 5 Total current of office from 17th hour to 24th hour**

For the 24 hours' current waveform in Fig.6.3 to Fig.6.5, there is no heater and refrigerator in office, so the current from the 1st hour to the 7th hour was unchanged since no one at the office. When an appliance is switched or the working state is changed, the total office current before switching is different from the current after

switching. The switching event is detected according to the change in total office current. After detecting the appliance switching event, the working current of the appliance is separated to identify the appliance type.

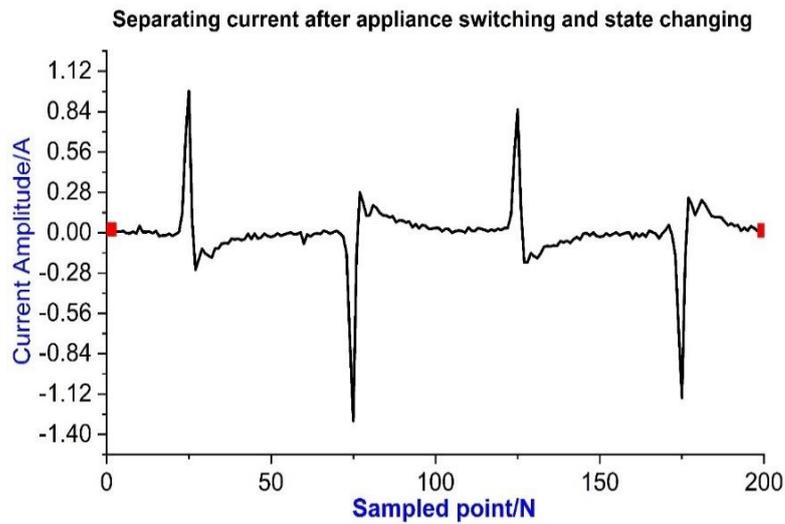
Taking one hour as an example for a detailed display, which was from 8:00 AM to 8:59 AM, Fig.6.6 shows the office's total current waveform.



**Fig.6. 6 The current waveform and current change point**

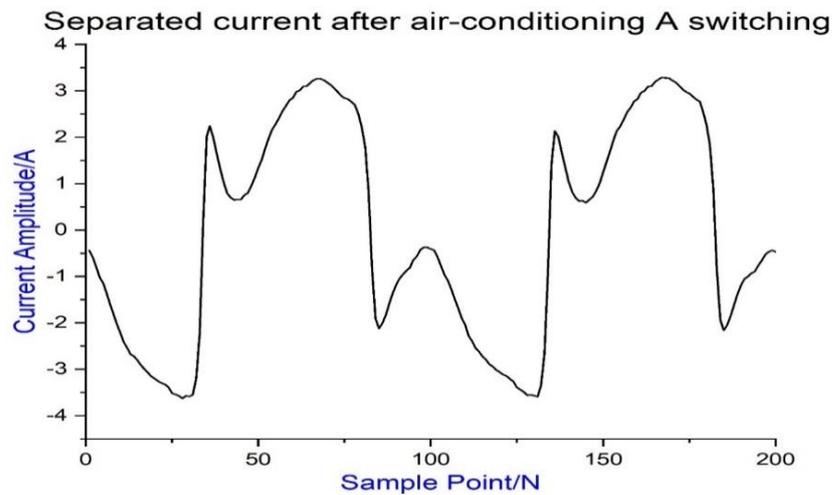
In Fig.6.6, the office's current waveform fluctuates and changes. According to the recorded data, the rms value of office current was calculated, if the rms value has changed, it indicates that some appliances switching or working state changed, then the appliance working current was separated and identified. The switching and state-changing points are marked with red lines in Fig.6.9, each switching and state-changing point will be described and analysed.

Firstly, the switching event was detected at the 2970000th sample point. The corresponding time point was 08:10:29. Then the working current of the switched appliance is separated, according to the appliance's working current frequency, using the KNN method and the appliance operation pattern to identify the separated current. The separated current waveform was identified as laptop-B. The result is shown in Fig.6.7.



**Fig.6. 7 Separated current after laptop-B switching**

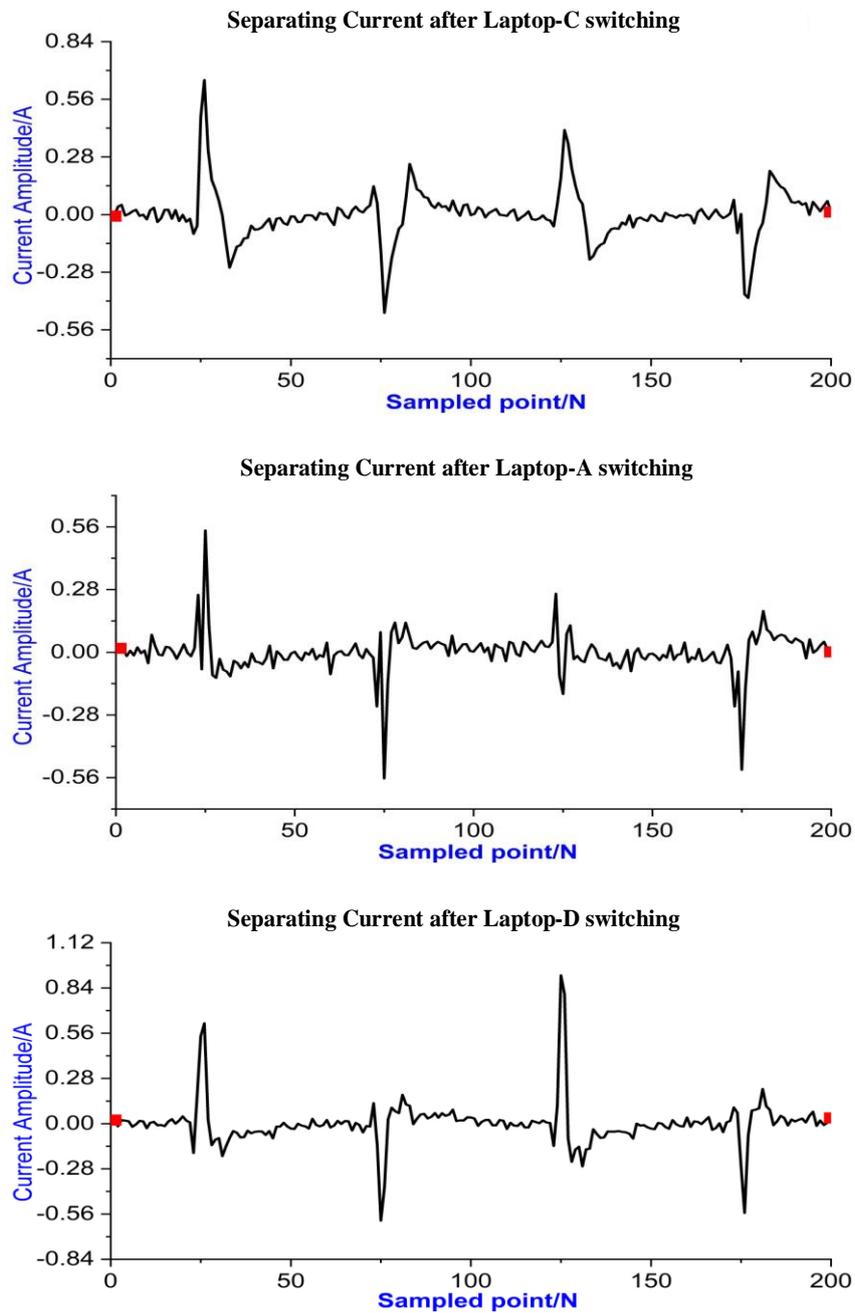
Secondly, the laptop-B operates stably after the switching event, so the total current is relatively stable. Then, the total office current increased greatly, the point of the switching event was detected at 4140000th sample point, and the corresponding time was 08:14:37. The working current of the switched appliance was separated, and the separated working current waveform was identified as air-conditioning A, the result is in Fig.6.8.



**Fig.6. 8 Separated current after air-conditioning A switching**

Thirdly, the office current reached the peak value after a period of time but did not increase instantaneously. This is because some appliances were turned on at almost the same time. Although some appliances were switched, each appliance reaches a

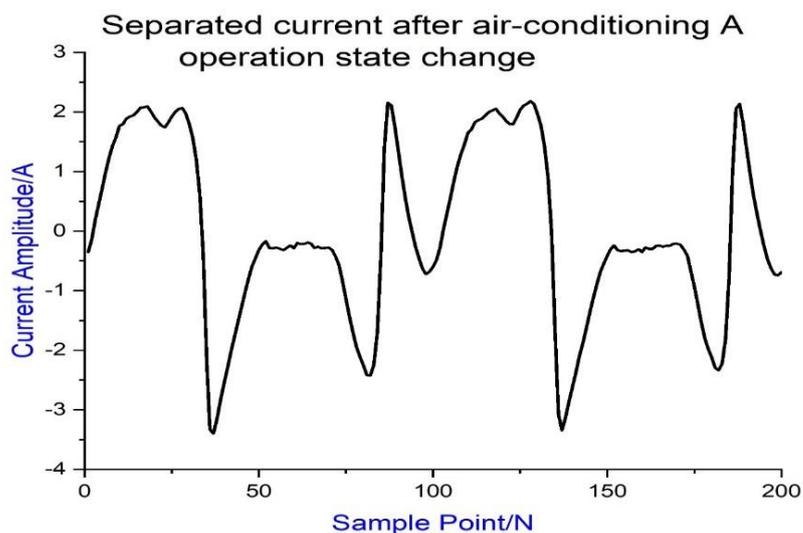
stable working state step by step. Therefore, three current waveforms were separated, the separated currents were identified as Laptop-C, Laptop-A, and Laptop-D respectively, and the corresponding switching time were 08:19:16, 08:20:39, and 08:23:18 respectively. The separated working current waveform of 3 laptops are shown in Fig.6.9.



**Fig.6. 9 Separated current after Laptop-C, Laptop-A and Laptop-D switching**

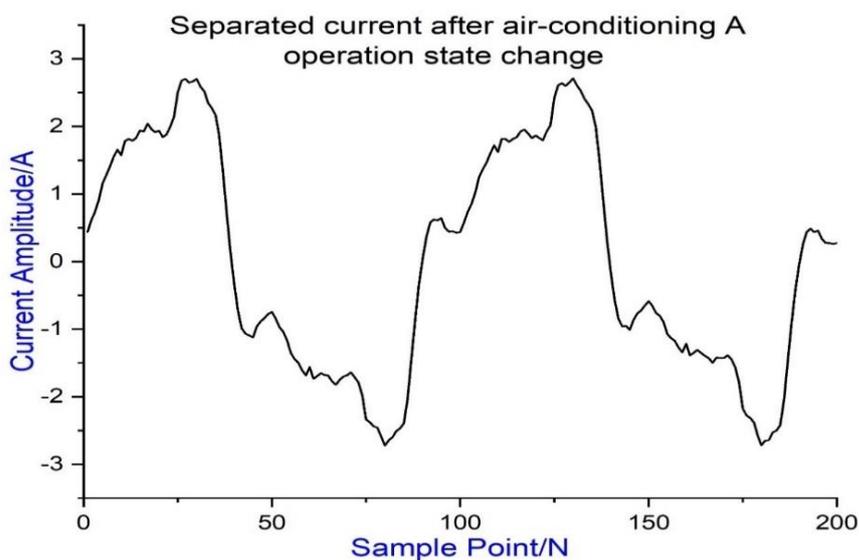
Fourthly, the total office current decreased greatly, which was caused by the air-

conditioning A operation state changes, the point of state-changing was detected at the 65100000th sample point which corresponds to the time point of 08:23:18. Then the working current of the appliance was separated and identified, the result is in Fig.6.10.



**Fig.6. 10 Separated current after air-conditioning A state change**

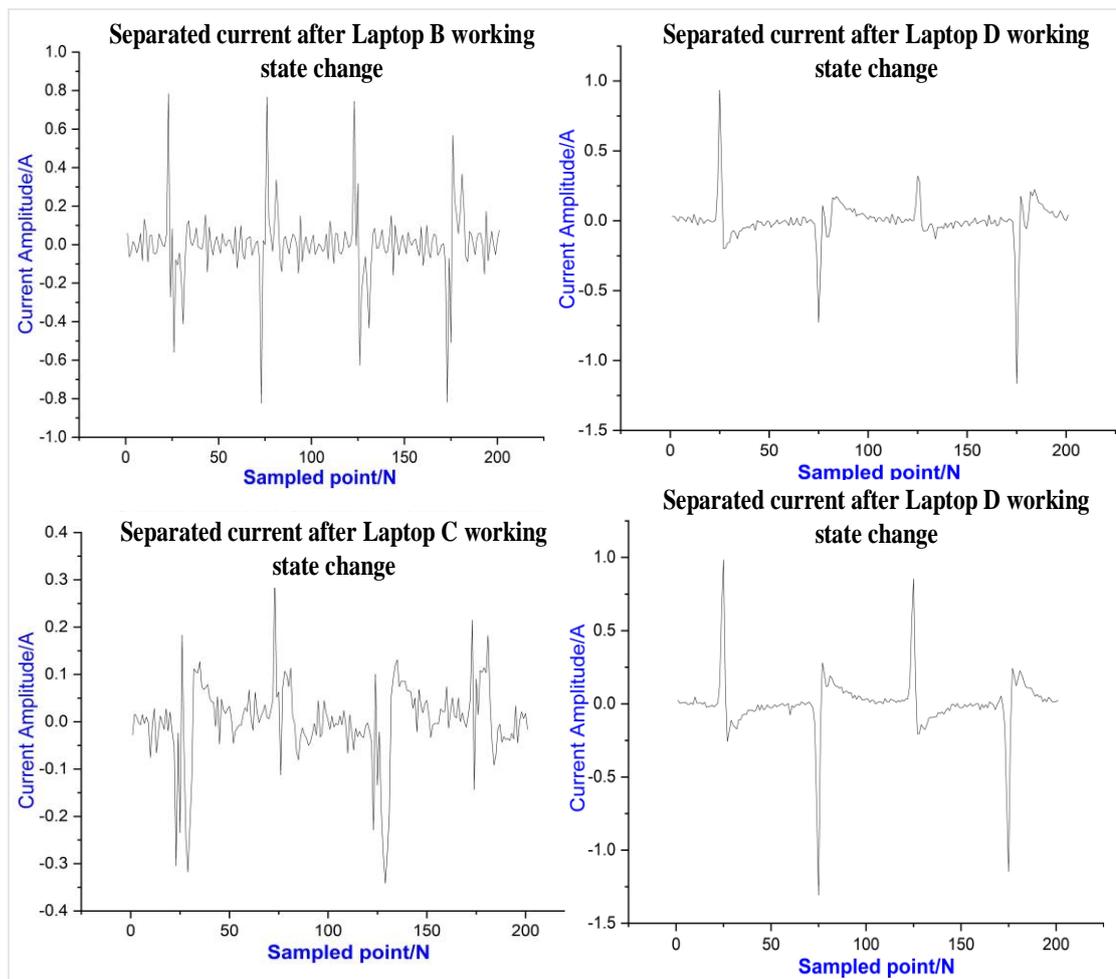
Fifthly, the operation state of the air-conditioning A changed again, so the office current increased. The state-changing was detected at the 74100000th sample point and the time was 08:26:09. The result is in Fig.6.11.



**Fig.6. 11 Separated current after air-conditioning A state change**

Finally, with four laptops and one air-conditioning working together in the office,

the office's current waveform fluctuated and changed continually, because the changing of the laptop's state is more frequent than other appliances. For example, when the laptop is editing documents and running the program MATLAB, the states of the CPU and other components are different, and the laptop's working current will be different under different states. Therefore, when four laptops work together, the office current looks like a noise current waveform, but the laptop's working state change can be detected, and the working current after state change can be separated. Four events about working state changes were detected, the result is shown in Fig.6.12.



**Fig.6. 12 Separated current after Laptop-B, Laptop-C and Laptop-D working state change**

Totally 232 switching events happened during the three days. Table 6.2 shows the switching point and corresponding time point.

**Table 6. 2 The appliance switching point and time point**

No.	Time Point						
1	08:10:29	59	17:48:23	117	09:27:13	175	01:21:11
2	08:14:37	60	17:48:30	118	09:34:31	176	01:21:17
3	08:18:43	61	17:48:36	119	09:53:16	177	09:38:07
4	08:19:16	62	17:50:18	120	11:32:56	178	09:38:13
5	08:20:39	63	17:50:24	121	11:36:00	179	09:40:20
6	08:23:18	64	17:54:13	122	11:36:06	180	09:40:27
7	08:23:24	65	17:59:11	123	11:45:57	181	09:40:46
8	08:26:09	66	18:19:00	124	11:46:04	182	09:40:52
9	08:27:32	67	18:20:26	125	12:20:07	183	09:40:52
10	08:29:58	68	18:42:40	126	12:24:21	184	09:41:24
11	08:30:04	69	18:58:46	127	12:24:28	185	09:45:25
12	08:31:01	70	19:21:04	128	12:31:33	186	09:54:19
13	10:00:32	71	19:25:18	129	12:31:40	187	09:54:32
14	10:11:00	72	19:41:37	130	12:35:47	188	09:56:52
15	10:07:44	73	19:41:43	131	12:40:14	189	09:57:30
16	11:34:44	74	19:41:49	132	12:55:35	190	10:03:04
17	11:36:00	75	19:42:02	133	14:10:23	191	10:05:49
18	11:36:57	76	19:42:15	134	14:21:49	192	10:06:15
19	11:37:04	77	19:55:42	135	14:22:01	193	10:07:44
20	11:37:10	78	19:56:52	136	14:44:03	194	10:29:13
21	11:37:23	79	19:57:11	137	14:44:09	195	10:35:03
22	11:38:02	80	20:21:36	138	14:44:09	196	10:43:05
23	11:38:26	81	20:31:01	139	14:44:16	197	12:40:27
24	11:38:32	82	20:31:40	140	14:44:22	198	12:45:32
25	11:40:08	83	20:31:46	141	14:55:54	199	13:06:28
26	11:40:46	84	20:31:52	142	14:56:32	200	13:14:43
27	11:42:21	85	20:32:56	143	14:58:14	201	13:58:46
28	11:47:26	86	20:37:42	144	16:13:04	202	14:21:42
29	11:48:36	87	20:51:21	145	18:02:01	203	14:43:25
30	11:48:42	88	20:51:28	146	18:31:46	204	15:05:11
31	11:58:14	89	20:53:09	147	18:32:11	205	15:05:18
32	12:26:41	90	20:53:22	148	18:32:43	206	15:06:02
33	12:41:05	91	22:12:36	149	18:35:41	207	15:06:08
34	12:43:37	92	22:23:18	150	18:42:15	208	15:11:20
35	13:15:40	93	22:23:56	151	18:43:50	209	15:21:04
36	13:20:02	94	22:24:15	152	18:48:04	210	15:22:46
37	13:30:36	95	22:24:21	153	20:20:39	211	15:22:52
38	13:31:08	96	22:49:14	154	20:55:54	212	15:39:36
39	13:31:14	97	22:50:18	155	20:56:45	213	15:39:42
40	13:42:40	98	23:43:50	156	20:56:52	214	15:40:40
41	13:42:53	99	00:10:48	157	20:56:52	215	15:41:11
42	13:55:54	100	00:10:54	158	20:57:04	216	15:43:50
43	13:56:01	101	00:12:04	159	21:02:01	217	15:43:56

**Table 6. 2 The appliance switching point and time point**

No.	Time Point						
44	13:56:07	102	00:27:57	160	21:10:10	218	15:58:20
45	13:56:20	103	08:59:18	161	21:17:09	219	16:06:21
46	13:56:26	104	08:59:24	162	21:17:41	220	16:30:55
47	14:09:32	105	09:38:00	163	21:19:42	221	16:33:28
48	14:09:38	106	09:44:00	164	21:56:32	222	19:14:49
49	14:09:57	107	09:09:13	165	21:59:18	223	19:15:08
50	14:10:04	108	09:12:36	166	22:16:06	224	19:33:02
51	14:10:16	109	09:12:55	167	22:16:12	225	19:33:28
52	14:10:35	110	09:13:14	168	22:33:40	226	19:46:23
53	14:10:42	111	09:13:27	169	22:35:09	227	19:52:12
54	14:11:26	112	09:13:27	170	22:36:06	228	19:52:18
55	14:28:04	113	09:16:12	171	01:13:40	229	20:41:49
56	17:47:20	114	09:18:00	172	01:15:08	230	20:42:02
57	17:47:32	115	09:21:04	173	01:15:21	231	20:47:07
58	17:47:39	116	09:21:23	174	01:21:04	232	22:22:08

According to the 232 times switching events, some events are caused by appliance operation state changing. For the working current of an appliance under different operating conditions, the frequency components of currents are similar since the similarity of main waveforms, and the main differences between currents are reflected in the amplitude, so according to the frequency features and current amplitude can know the appliance operation state changing. Furthermore, some switching events are happening at the same time, multiple appliances switching does not affect the current separation. The number of switching events for appliance operation state change, single appliance switching, and multiple appliance switching are shown in Table 6.3 individually.

**Table 6. 3 The Switching event classifying**

Appliance operation state changing times	Single appliance switching times	Multiple appliances switching times
171	37	21

Table 6.2 shows 232 switching events, but some appliances switching-off immediately after being switched on. Therefore, some switching events are filtered out, and the 229 switching events remain in Table 6.3.

When detecting appliance switching event, the working current of the appliance will be separated to identify the appliance type. The identification result of 229 switching events is in Table 6.4.

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
1	08:10:29	Laptop-B switch on	116	09:25:43	Laptop-B state change
2	08:14:37	Air-conditioning A switch on	117	09:27:13	Laptop-B state change
3	08:19:16	Laptop-C switch on	118	09:34:31	Laptop-C state change
4	08:20:39	Laptop-A switch on	119	09:53:16	Laptop-A state change
5	08:23:18	Laptop-D switch on	120	11:32:56	Kettle switch on
6	08:23:18	Air-conditioning A state change	121	11:36:00	Kettle switch off
7	08:26:09	Air-conditioning A state change	122	11:36:06	TV switch on
8	08:27:32	Laptop-B state change	123	11:44:57	Kettle switch on
9	08:29:58	Laptop-C state change	124	11:46:04	Kettle switch off
10	08:30:04	Laptop-D state change	125	12:20:07	Laptop-D state change
11	08:31:01	Laptop-D state change	126	12:24:21	Kettle switch on
12	10:00:32	Air-conditioning B switch on	127	12:24:28	Kettle switch off
13	10:11:00	Laptop-D state change	128	12:31:33	Laptop-A state change
14	10:07:44	Air-conditioning A state change	129	12:31:40	Laptop-A state change
15	11:34:44	Laptop-A state change	130	12:35:47	TV switch off
16	11:36:00	Air-conditioning A state change	131	12:40:14	Laptop-B switch off
17	11:36:57	Laptop-B state change	132	12:55:35	Laptop-A switch off
18	11:37:04	Laptop-A state change	133	14:10:23	Geyser switch on
19	11:37:10	Laptop-A state change	134	14:21:49	Laptop-B switch on

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
20	11:37:23	Laptop-D state change	135	14:22:01	Laptop-B state change
21	11:38:02	Air-conditioning A state change	136	14:44:03	Geysers switch off
22	11:38:26	Air-conditioning B state change	137	14:44:09	Laptop-A switch on
23	11:38:32	Air-conditioning B state change	138	14:44:09	TV switch on
24	11:40:08	Laptop-A state change	139	14:44:16	Laptop-D state change
25	11:40:46	Laptop-D state change	140	14:44:22	Laptop-C state change
26	11:42:21	Laptop-D state change	141	14:55:54	Geysers switch on
27	11:47:26	Laptop-C state change	142	14:56:32	Laptop-B state change
28	11:48:36	Air-conditioning A switch off	143	14:58:14	Laptop-A state change
29	11:48:42	Air-conditioning A switch off	144	16:13:04	Geysers switch off
30	11:58:14	Laptop-C state change	145	18:02:01	Laptop-B state change
31	12:26:41	Laptop-A switch off	146	18:31:46	Laptop-D state change
32	12:41:05	Laptop-B switch off	147	18:32:11	Laptop-A state change
33	12:43:37	Laptop-D switch off	148	18:32:43	Laptop-C state change
34	13:15:40	Laptop-C switch off	149	18:35:41	Laptop-B state change
35	13:20:02	Air-conditioning A switch on	150	18:42:15	Laptop-B state change
36	13:30:36	Air-conditioning A state change	151	18:43:50	Laptop-B state change
37	13:31:08	Air-conditioning B switch on	152	18:48:04	Laptop-A state change
38	13:31:14	Air-conditioning B state change	153	20:20:39	Laptop-A state change
39	13:42:40	Laptop-A switch on	154	20:55:54	Laptop-D state change
40	13:42:53	Laptop-A state change	155	20:56:45	Kettle switch on

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
41	13:55:54	Laptop-D switch on	156	20:56:52	Kettle switch off
42	13:56:01	Laptop-D state change	157	20:56:52	Laptop-B state change
43	13:56:07	Air-conditioning B state change	158	20:57:04	Laptop-A state change
44	13:56:20	Laptop-B switch on	159	21:02:01	Laptop-C state change
45	13:56:26	Laptop-B state change	160	21:10:10	Laptop-A state change
46	14:09:32	Laptop-C switch on	161	21:17:09	Laptop-D state change
47	14:09:38	Laptop-A state change	162	21:17:41	Laptop-A state change
48	14:09:57	Laptop-C state change	163	21:19:42	Laptop-D state change
49	14:10:04	Laptop-D state change	164	21:56:32	Laptop-A state change
50	14:10:16	Laptop-A state change	165	21:59:18	Laptop-C state change
51	14:10:35	Laptop-C state change	166	22:16:06	Laptop-D state change
52	14:10:42	Laptop-B state change	167	22:16:12	Laptop-B state change
53	14:11:26	Laptop-D state change	168	22:33:40	Laptop-A state change
54	14:28:04	Laptop-A state change	169	22:35:09	Laptop-A state change
55	17:47:20	Laptop-C state change	170	22:36:06	Laptop-B switch off
56	17:47:32	Laptop-D state change	171	01:13:40	Laptop-D switch off
57	17:47:39	Laptop-B state change	172	01:15:08	Laptop-A switch off
58	17:48:23	Laptop-A state change	173	01:15:21	Laptop-C switch off
59	17:48:30	Laptop-A state change	174	01:21:04	Unknown
60	17:48:36	Laptop-D state change	175	01:21:11	Unknown
61	17:50:18	Laptop-B state change	176	01:21:17	Unknown

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
62	17:50:24	Laptop-C state change	177	09:38:07	Laptop-C switch on
63	17:54:13	Laptop-A state change	178	09:38:13	Laptop-C state change
64	17:59:11	Laptop-C state change	179	09:40:20	Laptop-B switch on
65	18:19:00	Laptop-B state change	180	09:40:27	Laptop-B state change
66	18:20:26	Laptop-A state change	181	09:40:46	Air-conditioning A switch on
67	18:42:40	Laptop-C state change	182	09:40:52	Laptop-D switch on
68	18:58:46	Laptop-A state change	183	09:40:52	Air-conditioning B switch on
69	19:21:04	Laptop-D state change	184	09:41:24	Laptop-C state change
70	19:25:18	Laptop-B state change	185	09:45:25	Laptop-B state change
71	19:41:37	Laptop-A state change	186	09:54:19	Laptop-D state change
72	19:41:43	Laptop-C state change	187	09:54:32	Laptop-B state change
73	19:41:49	Laptop-B state change	188	09:56:52	Laptop-D state change
74	19:42:02	Laptop-D state change	189	09:57:30	Laptop-C state change
75	19:42:15	Laptop-A state change	190	10:03:04	Laptop-B state change
76	19:55:42	Laptop-C state change	191	10:05:49	Geyser switch on
77	19:56:52	Laptop-B state change	192	10:06:15	Laptop-A switch on
78	19:57:11	Laptop-B state change	193	10:07:44	Laptop-A state change
79	20:21:36	Laptop-A state change	194	10:29:13	Geyser switch off
80	20:31:01	Laptop-C state change	195	10:35:03	Microwave oven switch on
81	20:31:40	Laptop-A state change	196	10:43:05	Microwave oven switch off
82	20:31:46	Laptop-D state change	197	12:40:27	Laptop-B switch off

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
83	20:31:52	Laptop-A state change	198	12:45:32	Laptop-D switch off
84	20:32:56	Laptop-B state change	199	13:06:28	Laptop-A switch off
85	20:37:42	Laptop-B state change	200	13:14:43	Laptop-C switch off
86	20:51:21	Laptop-C state change	201	13:58:46	Geyser switch on
87	20:51:28	Laptop-A state change	202	14:21:42	Air-conditioning B switch off
88	20:53:09	Laptop-C state change	203	14:43:25	Geyser switch off
89	20:53:22	Laptop-D state change	204	15:05:11	Laptop-B state change
90	22:12:36	Laptop-B state change	205	15:05:18	Laptop-B state change
91	22:23:18	Laptop-A state change	206	15:06:02	Laptop-A state change
92	22:23:56	Laptop-C state change	207	15:06:08	Laptop-C state change
93	22:24:15	Laptop-D state change	208	15:11:20	Laptop-A state change
94	22:24:21	Laptop-A state change	209	15:21:04	Laptop-D state change
95	22:49:14	Laptop-B state change	210	15:22:46	Laptop-B state change
96	22:50:18	Air-conditioning B switch off	211	15:22:52	Laptop-C state change
97	23:43:50	Laptop-C switch off	212	15:39:36	Laptop-A state change
98	00:10:48	Laptop-D switch off	213	15:39:42	Laptop-C state change
99	00:10:54	Laptop-A switch off	214	15:40:40	Laptop-A state change
100	00:12:04	Laptop-B switch off	215	15:41:11	Laptop-C state change
101	00:27:57	Air-conditioning A switch off	216	15:43:50	Laptop-B state change
102	08:59:18	Laptop-C switch on	217	15:43:56	Laptop-A state change
103	08:59:24	Laptop-C state change	218	15:58:20	Laptop-C state change

**Table 6. 4 The identification result of 229 switching events(3days)**

<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>	<b>No.</b>	<b>Switching Time</b>	<b>Identifying Result</b>
104	09:38:00	Laptop-A switch on	219	16:06:21	Laptop-A state change
105	09:44:00	Laptop-B switch on	220	16:30:55	Laptop-D state change
106	09:09:13	Laptop-A state change	221	16:33:28	Air-conditioning B switch on
107	09:12:36	Laptop-D switch on	222	19:14:49	Laptop-B state change
108	09:12:55	Laptop-D state change	223	19:15:08	Laptop-A state change
109	09:13:14	Laptop-B state change	224	19:33:02	Air-conditioning A switch on
110	09:13:27	Geysler switch on	225	19:46:23	Laptop-A state change
111	09:13:27	Laptop-B state change	226	19:52:12	Laptop-B state change
112	09:16:12	Laptop-B state change	227	20:42:02	Laptop-A switch off
113	09:18:00	Laptop-C state change	228	20:47:07	Laptop-B switch off
114	09:21:04	Laptop-A state change	229	22:22:08	Air-conditioning A switch off
115	09:21:23	Geysler switch off			

After obtaining the appliance switching detection and identification result, the monitored results compare with actual switching conditions. The comparison is in Table 6.5.

**Table 6. 5 The monitored and real switching condition**

	<b>Real Switching times (Data)</b>	<b>Monitored Switching times (Test results)</b>
Appliance operation state changing times	195	171
Single appliance switching times	38	37
Multiple appliances switching times	29	21

According to Table 6.5, the appliance operation state changing occupies more than 80% of all switching events, and only 9% of events are multiple switching. 90% of switching events can be detected and identified through the proposed method.

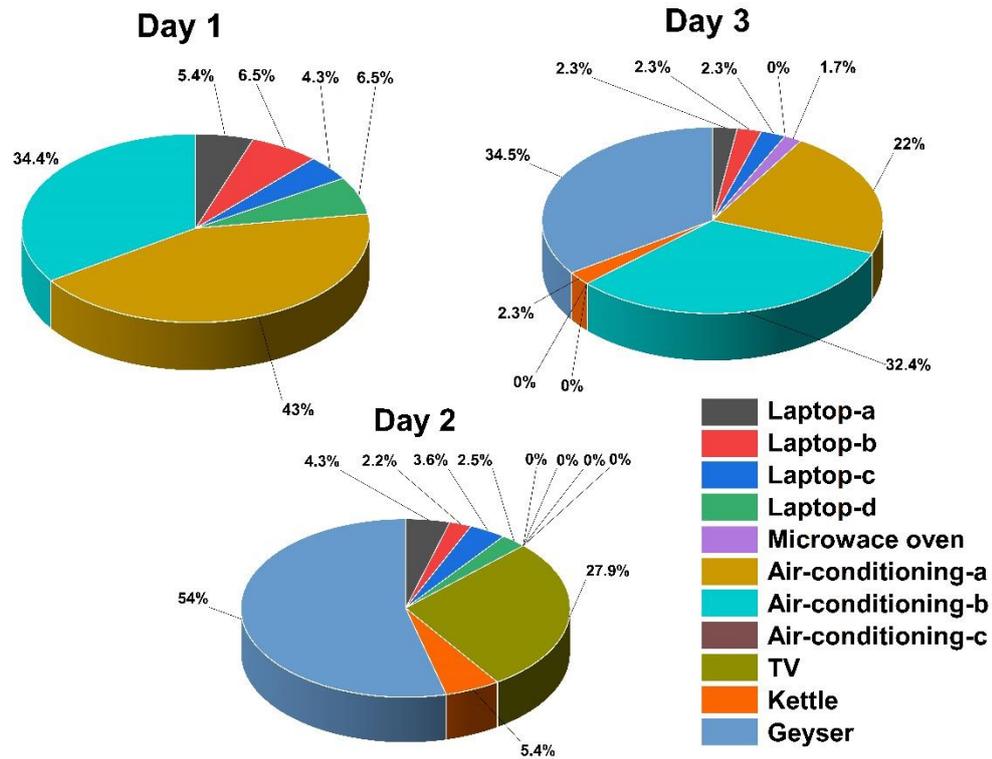
### 6.3.2 Power consumption monitoring

Based on the result in section 6.3.1, once knowing the appliance type, the appliance working current and voltage are already obtained, the appliance power can be calculated according to equation (3.8). Then according to the appliance switched ON and switched OFF time, the appliance power consumption can be calculated. The actual power consumption is also directly measured by the connected intelligent switchers which are used to check the calculated values. Table 6.6 shows the monitored power consumption of all appliances and the actual power consumption.

**Table 6. 6 The monitored power consumption**

Appliance Power Consumption (kWh)	Day 1		Day 2		Day 3	
	Monitored Value	Real Value	Monitored Value	Real Value	Monitored Value	Real Value
Laptop-A	0.103	0.1	0.135	0.12	0.101	0.11
Laptop-B	0.129	0.12	0.0717	0.06	0.102	0.11
Laptop-C	0.076	0.08	0.029	0.1	0.105	0.11
Laptop-D	0.126	0.12	0.078	0.07	0	0
Microwave oven	0	0	0	0	0.086	0.08
Air conditioning A	0.83	0.8	0	0	1.025	1.04
Air conditioning B	0.7	0.64	0	0	1.45	1.53
Air conditioning C	0	0	0	0	0	0
TV	0	0	0.723	0.77	0	0
Kettle	0	0	0.136	0.15	0.109	0.11
Geysers	0	0	1.37	1.49	1.421	1.63

When obtaining the power consumption of different appliances, the proportion of different loads in total energy consumption can be calculated, it is shown in Fig. 6.13.



**Fig.6. 13 Proportion of monitored electricity consumption in each day**

In Fig.6.13, the laptops account for a low proportion of power consumption among the users. Air conditioning is a high energy consumption appliance, and the proportion of power consumption by air conditioning is more than 30%; the geyser also consumes about 33% of total energy. Therefore, the operation of these two kinds of appliances in our daily life needs to decrease. Although the kettle also a high-power load, higher than a laptop, the operation period is shorter, so power consumption is relatively low.

#### **6.4 One Day On-line Test (24hours)**

This section is focused on the actual on-line testing of the proposed method, which involves a direct connection to the domestic homes. The duration of the tests is 24 hours (1day). The monitoring device was installed at the electricity supply point to capture the total current and voltage of these homes. The original data is then directly transmitted and processed, without any pre-processing of the data.

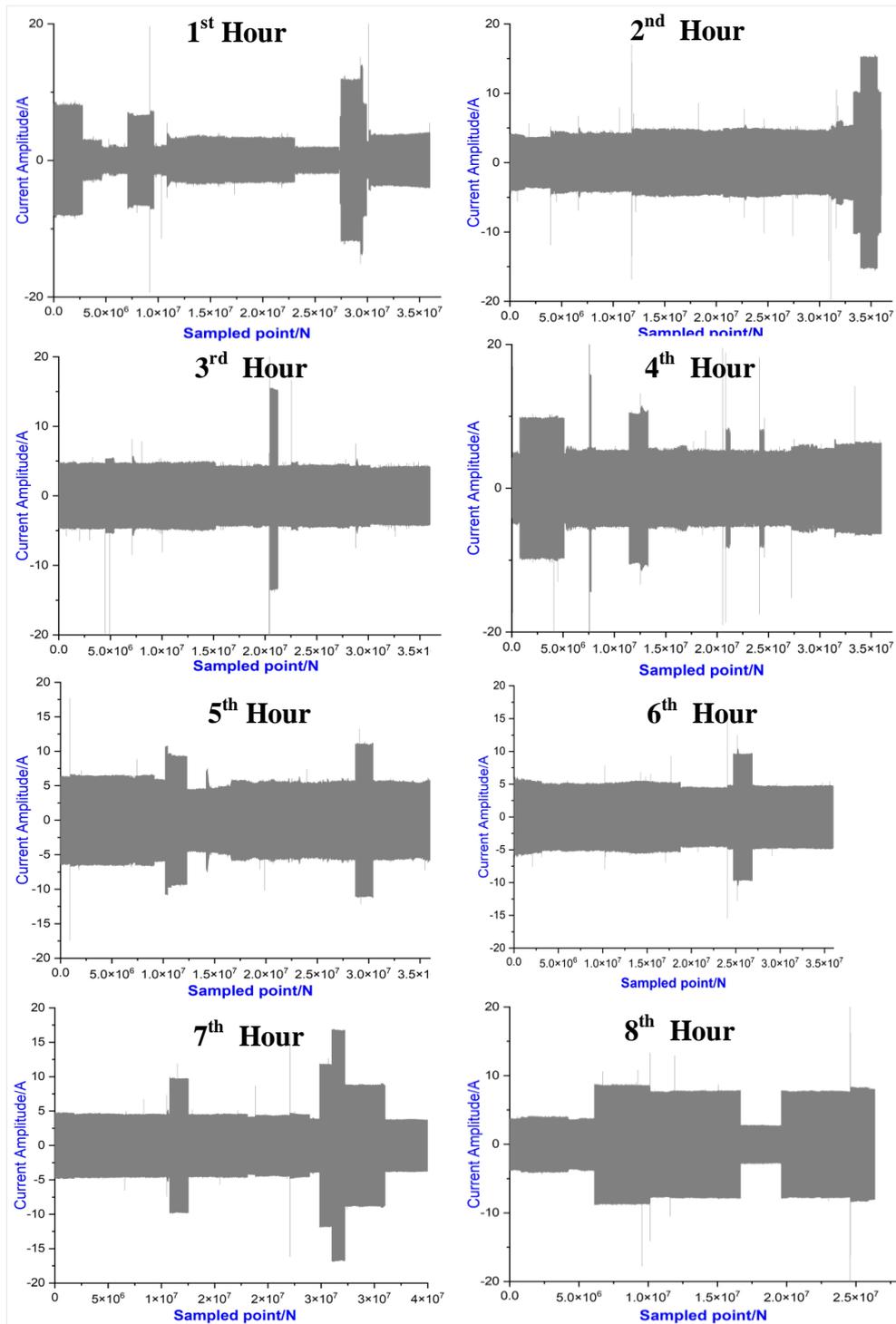
#### 6.4.1 Appliance switching detection and type identification

The proposed method was applied to 4 homes, which are named Home A, Home B, Home C, and Home D respectively. The main monitoring objects are shown in Table 6.7, these 6 appliances' power consumption is monitored respectively.

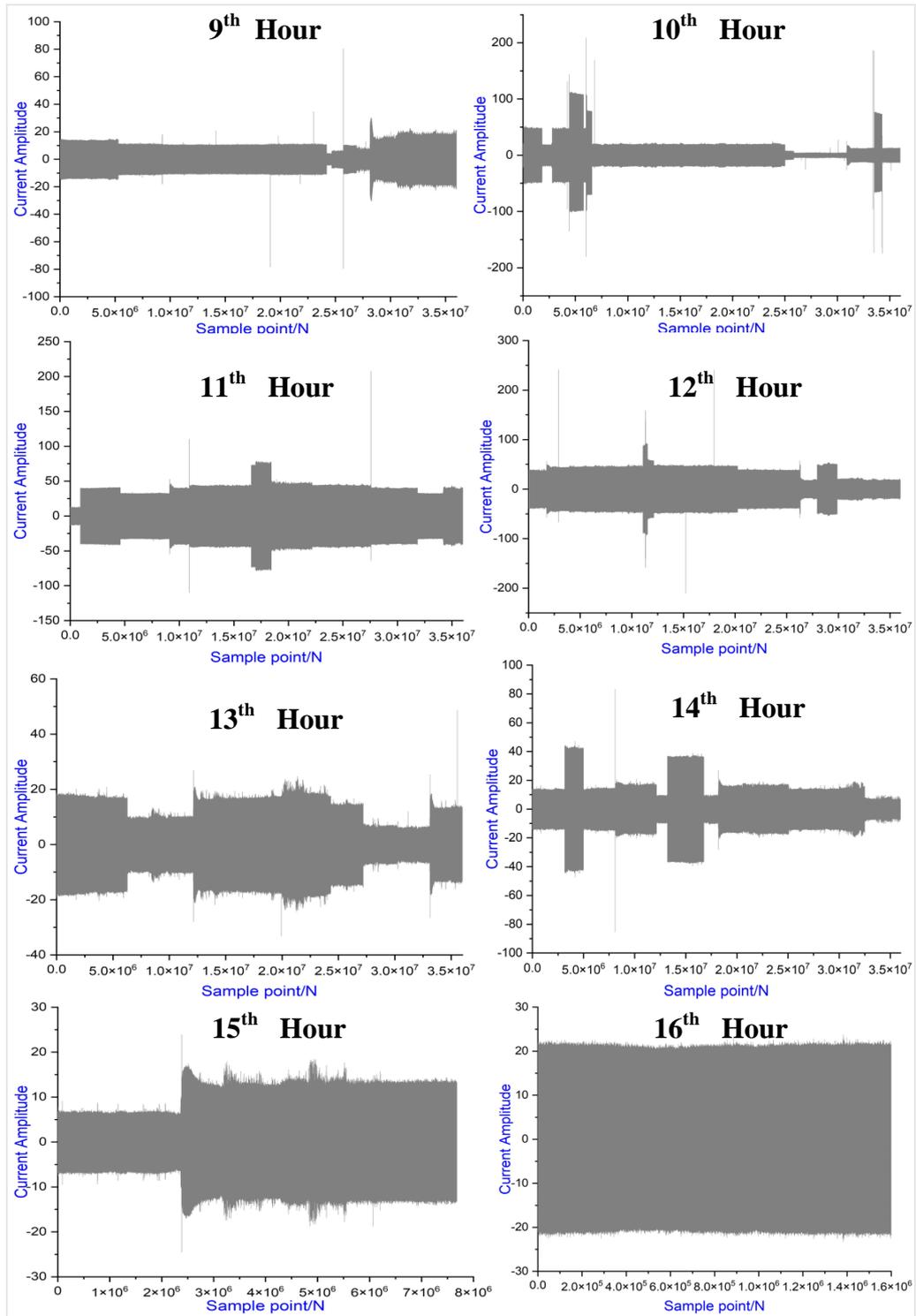
**Table 6. 7 Appliance name and type**

Appliance Name	Abbreviation for monitoring
Air-conditioning	AC
Laptop	LT
Kettle	EK
TV	TV
Geyser	WH
Refrigerator	RF

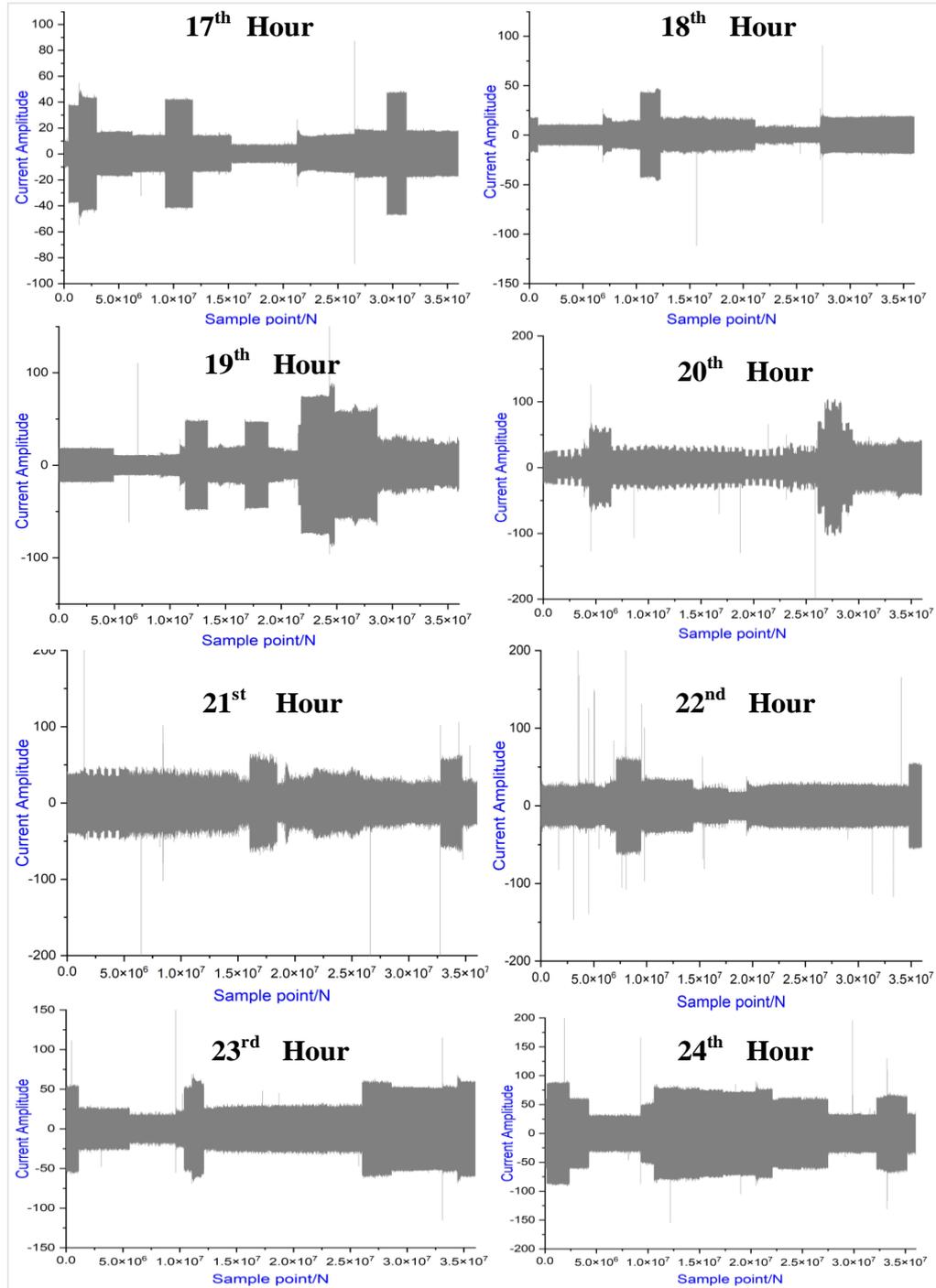
Taking home A as an example for a detailed display, during the monitoring process, the total home current was recorded and provided the reference data for identification results. Fig.6.14 to Fig.6.16 shows the 24 hours' current data of Home A, the data sample frequency is 10000Hz, so around  $3.6 \times 10^7$  data were recorded every hour. The current waveform of Homes B, C, and D are in Appendix II .



**Fig.6. 14**The 1st hour to 8th hour total current waveform of home A



**Fig.6. 15** The 12th hour to 16th hour total current waveform of home A



**Fig.6. 16 The 17th hour to 24th hour total current waveform of home A**

Based on the total home current, the switching event is detected by the change in total home current, the appliance switching detection results of Home A in 24 hours are shown in Table 6.8

**Table 6. 8 The appliance switching event and switching time point of Home A (24hours)**

<b>NO.</b>	<b>Time Point</b>						
<b>1</b>	16:04:42	<b>33</b>	22:20:56	<b>65</b>	04:45:02	<b>97</b>	10:31:26
<b>2</b>	16:07:45	<b>34</b>	22:20:58	<b>66</b>	04:55:16	<b>98</b>	10:35:56
<b>3</b>	16:11:51	<b>35</b>	22:41:26	<b>67</b>	05:05:18	<b>99</b>	10:35:58
<b>4</b>	16:16:03	<b>36</b>	22:43:22	<b>68</b>	05:08:23	<b>100</b>	10:41:24
<b>5</b>	16:18:06	<b>37</b>	22:45:28	<b>69</b>	05:13:32	<b>101</b>	10:47:48
<b>6</b>	16:45:51	<b>38</b>	22:51:44	<b>70</b>	05:20:32	<b>102</b>	11:07:38
<b>7</b>	16:49:21	<b>39</b>	23:10:12	<b>71</b>	05:22:02	<b>103</b>	11:10:08
<b>8</b>	16:50:15	<b>40</b>	23:27:05	<b>72</b>	05:27:58	<b>104</b>	11:43:32
<b>9</b>	17:55:33	<b>41</b>	23:32:42	<b>73</b>	05:30:22	<b>105</b>	11:44:38
<b>10</b>	17:56:42	<b>42</b>	0:40:22	<b>74</b>	05:41:05	<b>106</b>	11:47:02
<b>11</b>	17:59:24	<b>43</b>	00:41:01	<b>75</b>	05:54:12	<b>107</b>	11:49:02
<b>12</b>	17:59:58	<b>44</b>	00:42:54	<b>76</b>	06:04:01	<b>108</b>	12:26:44
<b>13</b>	18:34:08	<b>45</b>	00:46:56	<b>77</b>	08:00:44	<b>109</b>	12:30:48
<b>14</b>	18:35:03	<b>46</b>	01:03:04	<b>78</b>	08:05:00	<b>110</b>	12:54:06
<b>15</b>	19:01:02	<b>47</b>	01:04:38	<b>79</b>	08:10:26	<b>111</b>	12:57:05
<b>16</b>	19:08:34	<b>48</b>	01:07:24	<b>80</b>	08:15:24	<b>112</b>	13:11:54
<b>17</b>	19:12:04	<b>49</b>	01:09:40	<b>81</b>	08:19:36	<b>113</b>	13:15:48
<b>18</b>	19:12:58	<b>50</b>	01:10:12	<b>82</b>	08:25:03	<b>114</b>	13:58:00
<b>19</b>	19:19:13	<b>51</b>	01:11:02	<b>83</b>	08:35:28	<b>115</b>	14:01:56
<b>20</b>	19:19:04	<b>52</b>	01:55:42	<b>84</b>	08:44:16	<b>116</b>	14:17:00
<b>21</b>	19:22:12	<b>53</b>	01:57:01	<b>85</b>	08:49:08	<b>117</b>	14:20:14
<b>22</b>	19:34:46	<b>54</b>	02:01:46	<b>86</b>	08:52:01	<b>118</b>	14:43:26
<b>23</b>	19:35:32	<b>55</b>	02:27:42	<b>87</b>	09:11:03	<b>119</b>	14:55:14
<b>24</b>	19:40:14	<b>56</b>	02:30:46	<b>88</b>	09:12:52	<b>120</b>	15:00:02
<b>25</b>	19:41:04	<b>57</b>	03:18:03	<b>89</b>	09:17:18	<b>121</b>	15:04:00
<b>26</b>	20:17:02	<b>58</b>	03:19:16	<b>90</b>	09:20:28	<b>122</b>	15:07:08
<b>27</b>	20:20:36	<b>59</b>	03:43:55	<b>91</b>	09:35:16	<b>123</b>	15:15:03
<b>28</b>	20:47:52	<b>60</b>	03:46:38	<b>92</b>	09:45:44	<b>124</b>	15:17:43
<b>29</b>	20:50:48	<b>61</b>	03:49:05	<b>93</b>	10:08:16	<b>125</b>	15:36:05
<b>30</b>	21:41:12	<b>62</b>	04:10:03	<b>94</b>	10:18:58	<b>126</b>	15:45:05
<b>31</b>	21:44:05	<b>63</b>	04:20:18	<b>95</b>	10:22:22	<b>127</b>	15:53:42
<b>32</b>	22:18:00	<b>64</b>	04:40:34	<b>96</b>	04:45:02	<b>128</b>	15:58:04

Totally 128 switching events are detected in Home A during 24 hours, the working current of a switched appliance is separated to identify the appliance type when detecting a switching event, based on the frequency of separate currents and the appliance operation pattern, and the identification result of 128 switching events are listed in Table 6.9.

**Table 6. 9 The identification result of 128 switching events in Home A**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
1	16:04:42	TV switch off	65	04:45:02	Refrigerator switch on
2	16:07:45	Laptop switch off	66	04:55:16	Refrigerator switch off
3	16:11:51	Refrigerator switch on	67	05:05:18	Geyser switch on
4	16:16:03	Unknown	68	05:08:23	Geyser switch off
5	16:18:06	Refrigerator switch off	69	05:13:32	Unknown
6	16:45:51	Refrigerator switch on	70	05:20:32	Unknown
7	16:49:21	Unknown	71	05:22:02	Refrigerator switch on
8	16:50:15	Refrigerator switch off	72	05:27:58	Refrigerator switch off
9	17:55:33	Refrigerator switch on	73	05:30:22	Unknown
10	17:56:42	Microwave oven switch on	74	05:41:05	Unknown
11	17:59:24	Microwave oven switch off	75	05:54:12	Unknown
12	17:59:58	Refrigerator switch off	76	06:04:01	Geyser switch on
13	18:34:08	Kettle switch on	77	08:00:44	Refrigerator switch on
14	18:35:03	Kettle switch off	78	08:05:00	Refrigerator switch off
15	19:01:02	Refrigerator switch on	79	08:10:26	Microwave oven switch on
16	19:08:34	Refrigerator switch off	80	08:15:24	Microwave oven switch off
17	19:12:04	Unknown	81	08:19:36	Laptop switch on
18	19:12:58	Unknown	82	08:25:03	Unknown
19	19:19:04	Noise	83	08:35:28	Unknown
20	19:19:13	Geyser switch on	84	08:44:16	Refrigerator switch on
21	19:22:12	Geyser switch off	85	08:49:08	Refrigerator switch off
22	19:34:46	Unknown	86	08:52:01	TV switch on
23	19:35:32	Refrigerator switch on	87	09:11:03	Kettle switch on
24	19:40:14	Unknown	88	09:12:52	Kettle switch off

**Table 6. 9 The identification result of 128 switching events in Home A**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
25	19:41:04	Refrigerator switch off	89	09:17:18	Refrigerator switch on
26	20:17:02	TV switch on	90	09:20:28	Refrigerator switch off
27	20:20:36	Air-conditioning switch on	91	09:35:16	Laptop switch off
28	20:47:52	Refrigerator switch on	92	09:45:44	Laptop switch on
29	20:50:48	Refrigerator switch off	93	10:08:16	Laptop switch off
30	21:41:12	Refrigerator switch on	94	10:18:58	Refrigerator switch on
31	21:44:05	Refrigerator switch off	95	10:22:22	Refrigerator switch off
32	22:18:00	Refrigerator switch on	96	10:27:02	Geyser switch on
33	22:20:56	Refrigerator switch off	97	10:31:26	Geyser switch off
34	22:20:58	Noise	98	10:35:56	Air-conditioning switch on
35	22:41:26	Geyser switch on	99	10:35:58	Air-conditioning switch on
36	22:43:22	Geyser switch off	100	10:41:24	Unknown
37	22:45:28	Air-conditioning switch off	101	10:47:48	TV switch off
38	22:51:44	TV switch off	102	11:07:38	Refrigerator switch on
39	23:10:12	Laptop switch on	103	11:10:08	Refrigerator switch off
40	23:27:05	Laptop switch off	104	11:43:32	Microwave oven switch on
41	23:32:42	Unknown	105	11:44:38	Kettle switch on
42	0:40:22	Laptop switch on	106	11:47:02	Kettle switch off
43	00:41:01	Unknown	107	11:49:02	Microwave oven switch off
44	00:42:54	Unknown	108	12:26:44	Refrigerator switch on
45	00:46:56	Geyser switch on	109	12:30:48	Refrigerator switch off
46	01:03:04	Geyser switch off	110	12:54:06	Geyser switch on

**Table 6. 9 The identification result of 128 switching events in Home A**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
47	01:04:38	Refrigerator switch on	111	12:57:05	Geysers switch off
48	01:07:24	Geysers switch on	112	13:11:54	Refrigerator switch on
49	01:09:40	Noise	113	13:15:48	Refrigerator switch off
50	01:10:12	Geysers switch off	114	13:58:00	TV switch on
51	01:11:02	Refrigerator switch off	115	14:01:56	Air-conditioning switch off
52	01:55:42	Refrigerator switch on	116	14:17:00	Refrigerator switch on
53	01:57:01	Refrigerator switch off	117	14:20:14	Refrigerator switch off
54	02:01:46	Unknown	118	14:43:26	Air-conditioning switch on
55	02:27:42	Refrigerator switch on	119	14:55:14	Laptop switch on
56	02:30:46	Refrigerator switch off	120	15:00:02	Refrigerator switch on
57	03:18:03	Refrigerator switch on	121	15:04:00	Refrigerator switch off
58	03:19:16	Refrigerator switch off	122	15:07:08	TV switch off
59	03:43:55	Geysers switch on	123	15:15:03	Refrigerator switch on
60	03:46:38	Refrigerator switch on	124	15:17:43	Refrigerator switch off
61	03:49:05	Geysers switch off	125	15:36:05	Laptop switch off
62	04:10:03	Refrigerator switch off	126	15:45:05	Air-conditioning switch off
63	04:20:18	Air-conditioning switch on	127	15:53:42	Refrigerator switch on
64	04:40:34	Air-conditioning switch off	128	15:58:04	Refrigerator switch off

After obtaining the monitoring result, the monitored result compares with the real switching condition, which is shown in Table 6.10.

**Table 6. 10 The monitored and real switching condition in Home A**

	<b>Real Switching times</b>	<b>Monitored Switching times</b>
Appliance operation state changing times	12	10
Single appliance switching times	129	114
Multiple appliances switching times	6	4

According to Table 6.10, in the domestic Home A, the single appliance switching occupies more than 80% of all switching events, but only 4 events were multiple appliances switched at the same time. Most switching events could be detected and identified through the proposed method.

121, 103, and 107 switching events are detected in Homes B, C, and D respectively during 24 hours, and their identification results are listed in Appendix III.

#### **6.4.2. Power consumption monitoring**

Through 24-hour monitoring, the appliance switching and identifying of Homes A, B, C, and D are obtained. Once knowing the appliance type, the appliance's working current and voltage are already sampled, based on equation (3.8), the appliance power can be obtained, then according to the appliance switched ON and switched OFF time, the appliance power consumption can be calculated. The monitored power consumption of 4 homes is in Table 6.11.

**Table 6. 11 Power consumption of appliances in four homes**

Appliance Power Consumption (kWh)		Appliance Name					
		AC	TV	LP	WH	EK	RF
Home A	Monitoring Value	44.005	4.619	0.174	8.440	10.908	13.755
Home B	Monitoring Value	55.018	1.611	0.106	6.269	17.992	16.123
Home C	Monitoring Value	41.619	4.412	0.198	11.519	17.799	24.76
Home D	Monitoring Value	24.976	8.472	0.397	11.136	34.912	15.936

Table 6.11 compares the power information of six monitored load power consumption within four different homes based on the proposed method. Based on power consumption, laptops account for a low proportion of power consumption among users. Air conditioning is a high energy consumption appliance, and the proportion of power consumption by air conditioning is more than 30% in three homes; the geyser also consumes about 31% of total energy in home 4. Therefore, the operation of these two kinds of loads in our daily life needs to decrease total energy demand. Furthermore, though the refrigerator is a low power consumption load, contrary to popular opinion, the operation period is long, power consumption is relatively high ranging from 5% to 12% depending on home usage.

### 6.5 Three Months Monitoring of Power Consumption for a Normal Home

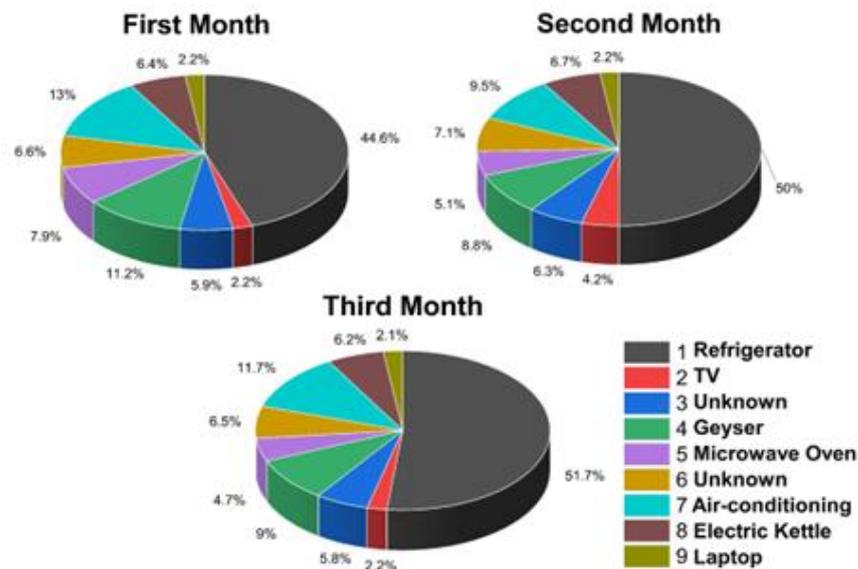
To verify the effectiveness of the proposed identification method over a longer time period, the power of a normal home was monitored for 90 days. The main appliances monitored were air conditioning (AC), TV, laptop (LT), electric kettle (EK), geyser (WH), refrigerator (RF) and microwave oven (MO). Besides these seven appliances, all appliances' switching events were also detected and tracked over 90 days. There were 35 operation state changing, 1173 single switching and 25 multiple switching events. Table 6.12 shows the monitored power consumption of nine appliances.

**Table 6. 12 Appliance power consumption**

<b>Appliance Name</b>	<b>Power (kW)</b>	<b>Operation Period(h)</b>	<b>Power Proportion(%)</b>
Microwave Oven	1.73	29.2	5.3%
Air Conditioning	2.125	176	12.5%
Electric Kettle	2.11	15.1	3.2%
Geyser	2.45	50.7	11.3%
Laptop	0.27	310	5.3%
Refrigerator	0.93	721	40.2%
TV	0.35	348	8.7%
Electric Kettle	2.11	15.1	3.2%
Unknown 1	0.55	—	9.1%
Unknown 2	0.53	—	4.1%

In Table 6.12, two of them were unidentified and marked as unknown appliances as there were no training samples in the KNN model that matched their appliance features. However, the power consumptions of these two unknown appliances are quite significant in the total power consumption, hence it is necessary to gather statistics on them to account for their power consumption compared to the other appliances. The power consumption of the unknown appliance can also be calculated using equation (3.8), which is based on the separation of current and voltage.

Fig.6.17 shows the calculated proportion of different loads in total energy consumption, providing information about the power consumption of different appliances. For example, the refrigerator is a high energy consumption appliance, contributing to a power proportion of over 40%. The air-conditioning also consumed about 12.5% of the total energy. Therefore, it is important to reduce the operations of these two kinds of loads in our daily life. Contrary to public opinion, the kettle is also a high-power consumption load. Although the operation period is often short, its power consumption can be relatively high.



**Fig.6. 17 Appliance power consumption proportion**

#### 6.4 Summary

This chapter provides the application and testing of the proposed model. The monitoring of household appliance running condition and power consumption is

achieved, which can improve power network stability for demand-side management.

The monitoring hardware is described first. To verify the effectiveness of monitoring, offline and online testing was performed. In the offline test, 3 days of monitoring were processed in a university office, which is based on the recorded data of office current and voltage. For an online test, the monitoring was processed in four domestic homes, 1 day's monitoring based on four households, and 90 days' monitoring based on one household. In both offline and online testing, the power consumption of each appliance was obtained.

## **Chapter 7 Conclusion and Future Work**

### **7.1 General conclusions**

This thesis focused on non-intrusive load monitoring (NILM) on residential sites with the primary goals of improving monitoring accuracy, decreasing monitoring response time, and enhancing the practicability of NILM. To achieve this, a physical monitoring device was constructed to capture live data. The methods used to analyse the captured data were continuously improved to achieve this project's aim.

Chapter 3 systematically compared and discussed the appliance features. Effective and valid features for appliance identification were extracted and sampled from real residential power environments. This prior research laid the foundation for further training and constructing of the monitoring model.

Chapter 4 improved on the detection method for appliance switching and the separation method for appliance working current, thereby reducing the response time for monitoring and enhancing the integrity of the separated features. This is done by improving the efficiency and sensitivity of the Heuristic detection method, and eliminating the use of the auxiliary time-meter during the detection process. Based on sampled data at the electricity supply point, the proposed method can detect any switching events of low or high-power appliances in real-time. Furthermore, the issue of the difference method not being able to obtain the phase angle of the separated current was resolved. The improved difference method separated the appliance working current with accurate phase angles from the mixed current sequence at the electricity supply point. Finally, the proposed detecting and separating method were tested using all the examples and cases from actual residential power environments, verifying the practicality of the method.

Chapter 5 proposed the identification method for the appliance type using the KNN method to reduce the response time and improve the identifying accuracy. First, the KNN model was reconstructed and trained according to the appliance feature analysis discussed in Chapter 3. The KNN model structure was simplified to shorten the data processing time, thereby reducing the response time and the storage space used to save training samples. Next, to address the issue of incorrect identification due

to similar electrical features in two different appliances, appliance operating patterns were used as auxiliary references for the appliance type identification. A learning strategy for obtaining the general operation pattern was proposed. The obtained operation patterns were verified and tested using real data. The results demonstrate that the operation pattern curve provides regularity and possible future trajectory of a running appliance. It is hence used to modify the identification result, which overcomes the problem of overlapping appliance electric features and improves the monitoring accuracy.

Chapter 6 tested and applied collectively the data acquisition device, the switching event detecting method, the appliance current separating method and the appliance type identifying method as a complete monitoring process. This process was applied both off-line and on-line. There were 232 switching events in the off-line test and 459 switching events in the on-line test. Upon detecting the switching event, the operation of each appliance was differentiated. The switching on-off time, frequency and duration of each operation were monitored, and the power consumption of each appliance during their operation period was obtained.

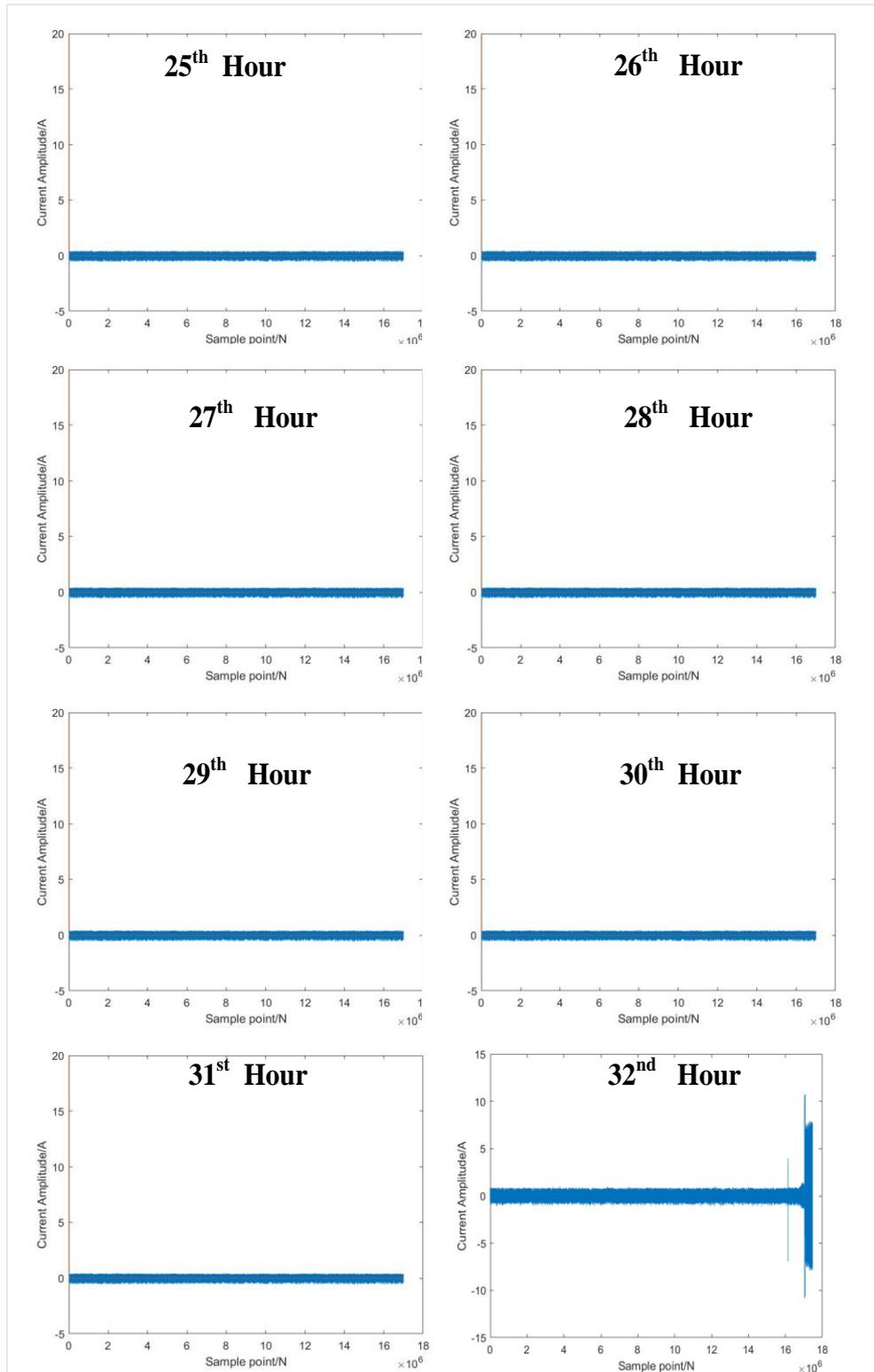
## **7.2 Suggestions for future research**

The research on the implementation and application of NILM presented in this thesis is non-exhaustive. Future work is recommended to address the following aspects:

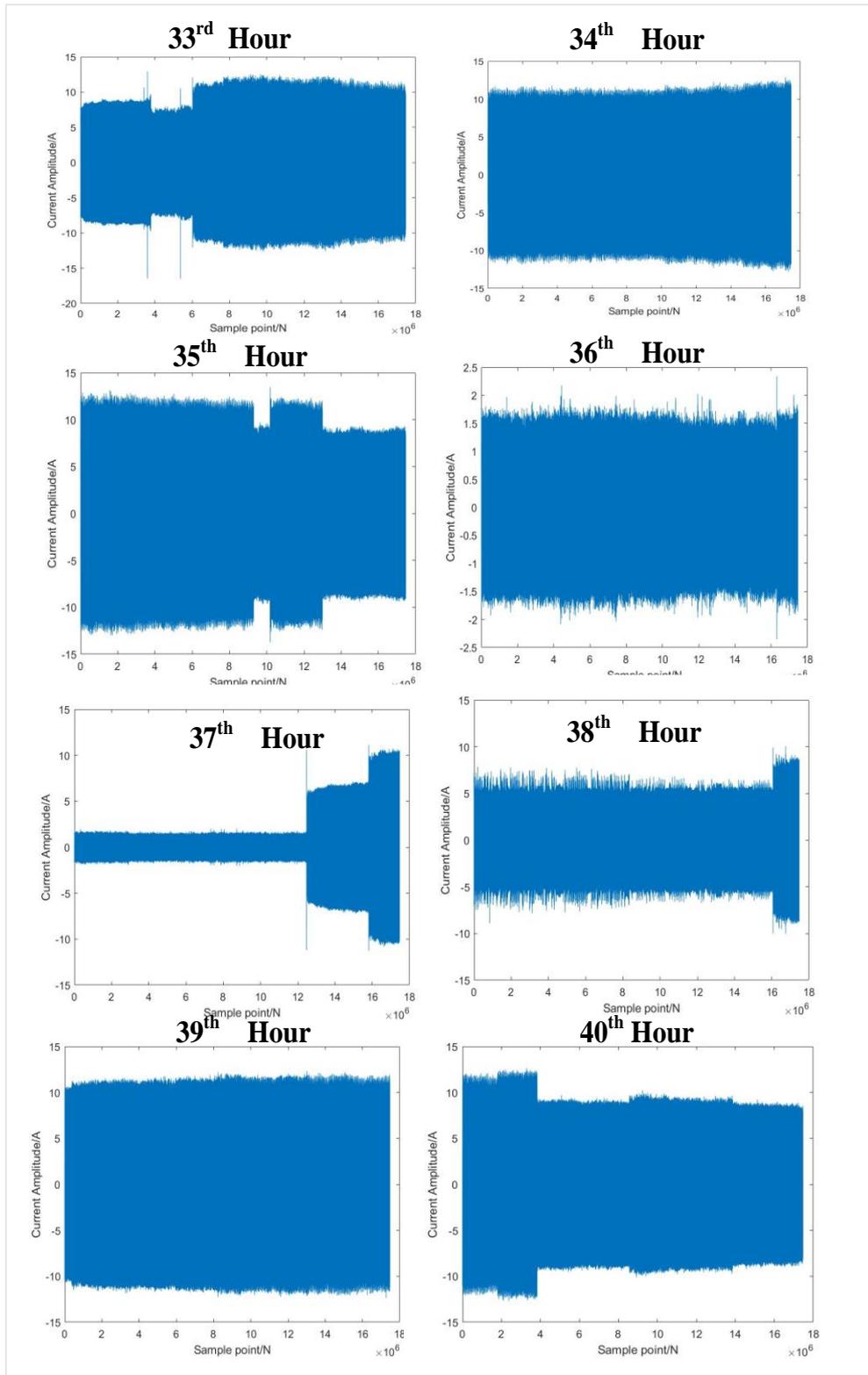
1. Parameters involved in the detection of the switching events are often predetermined based on prior analysis, such as setting the observation period as 2 seconds. This can affect convergence speed, separation accuracy, and application promotion of the identification algorithm. Therefore, future research with an adaptive parameter selection strategy is suggested.
2. In this thesis, the appliance's electric and non-electric features were extracted manually. However, the electric features of different appliances always have subtle differences, and non-electric features also require updates according to the seasons and years. Therefore, realizing the self-extraction of household appliance features is crucial for the further development of NILM.
3. In the KNN identification process, as there are only one or two home appliances belonging to the same category, the number of training samples per

category is therefore very few. This thesis proposes giving different weightings to each category in the majority identification process. However, this may result in a decrease in the identification accuracy. Future work can analyse more appliances to enlarge the case number in each category and treat all these analysed appliances as training samples for the KNN process.

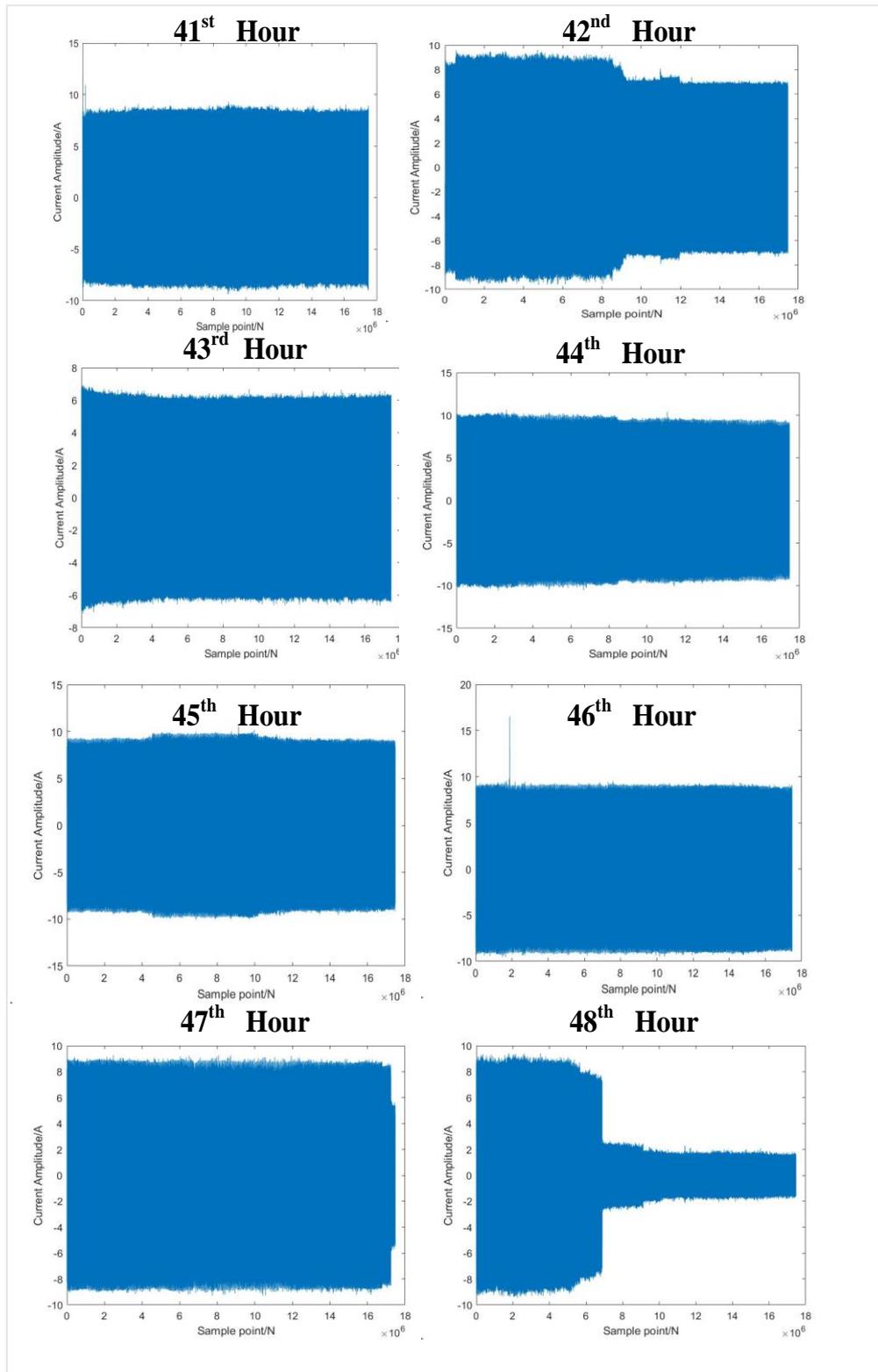
## Appendix I Data: Office current waveform from 25th hour to 72nd hour



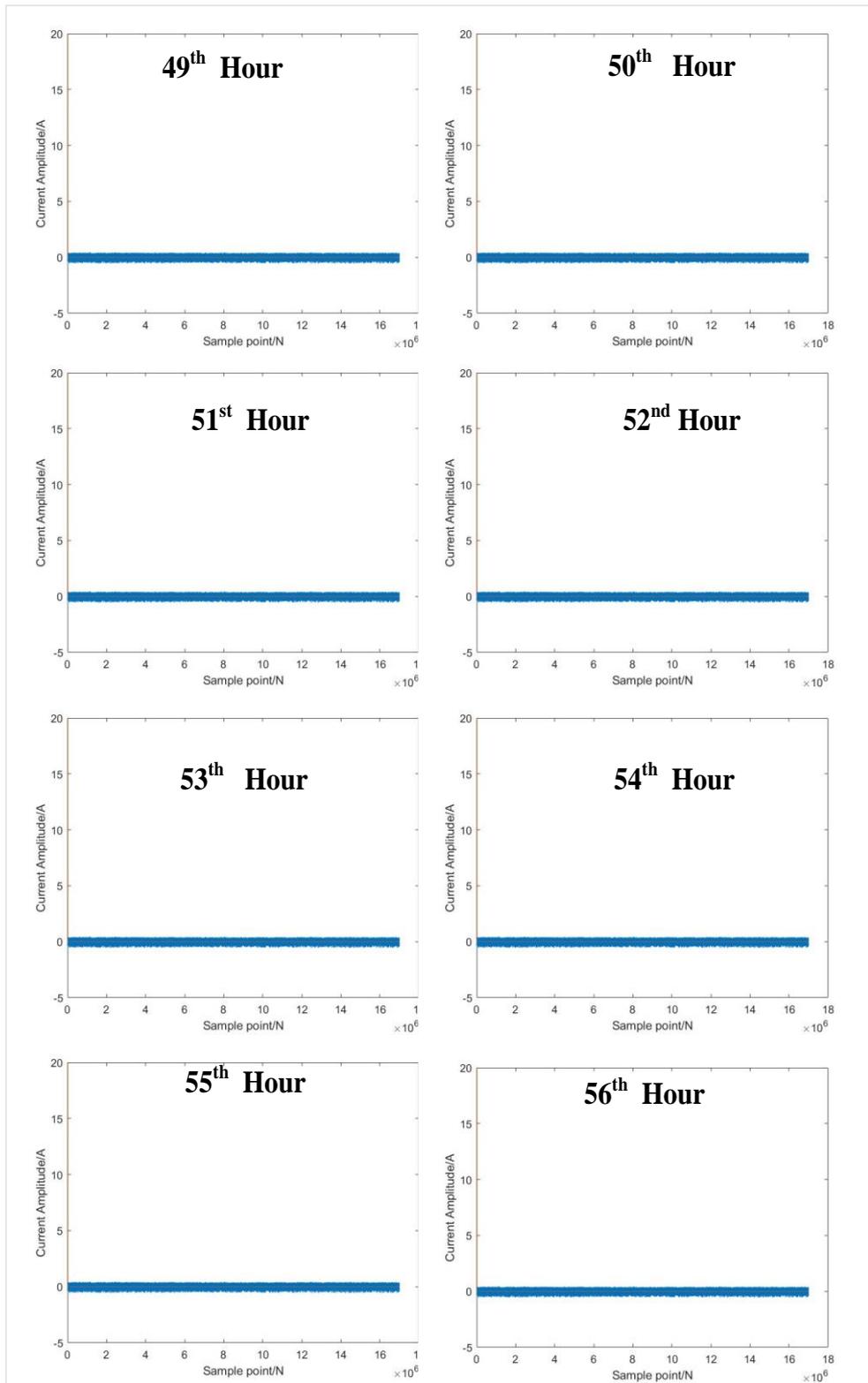
Total current of office from 25<sup>th</sup> hour to 32<sup>nd</sup> hour



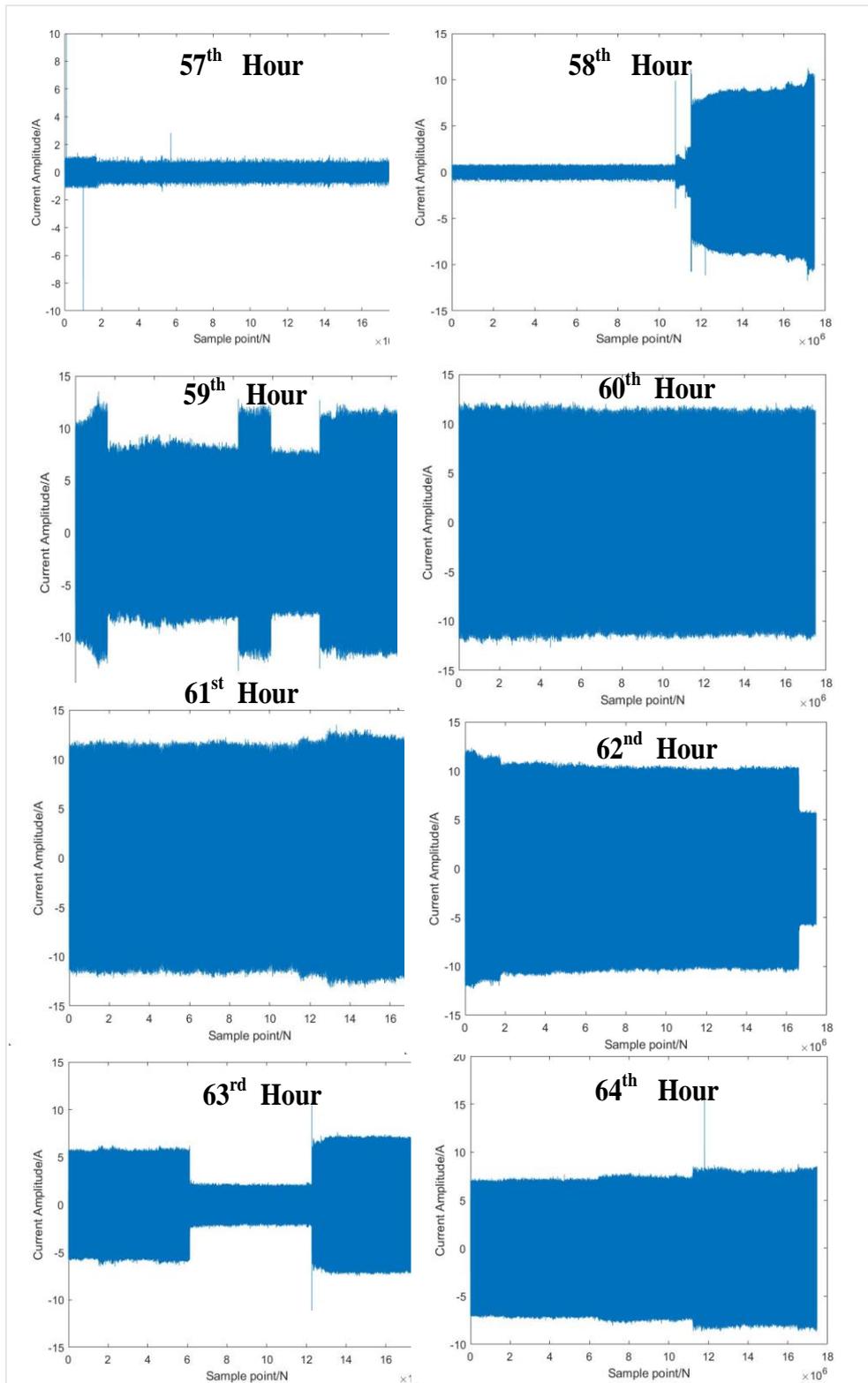
**Total current of office from 33<sup>rd</sup> hour to 40<sup>th</sup> hour**



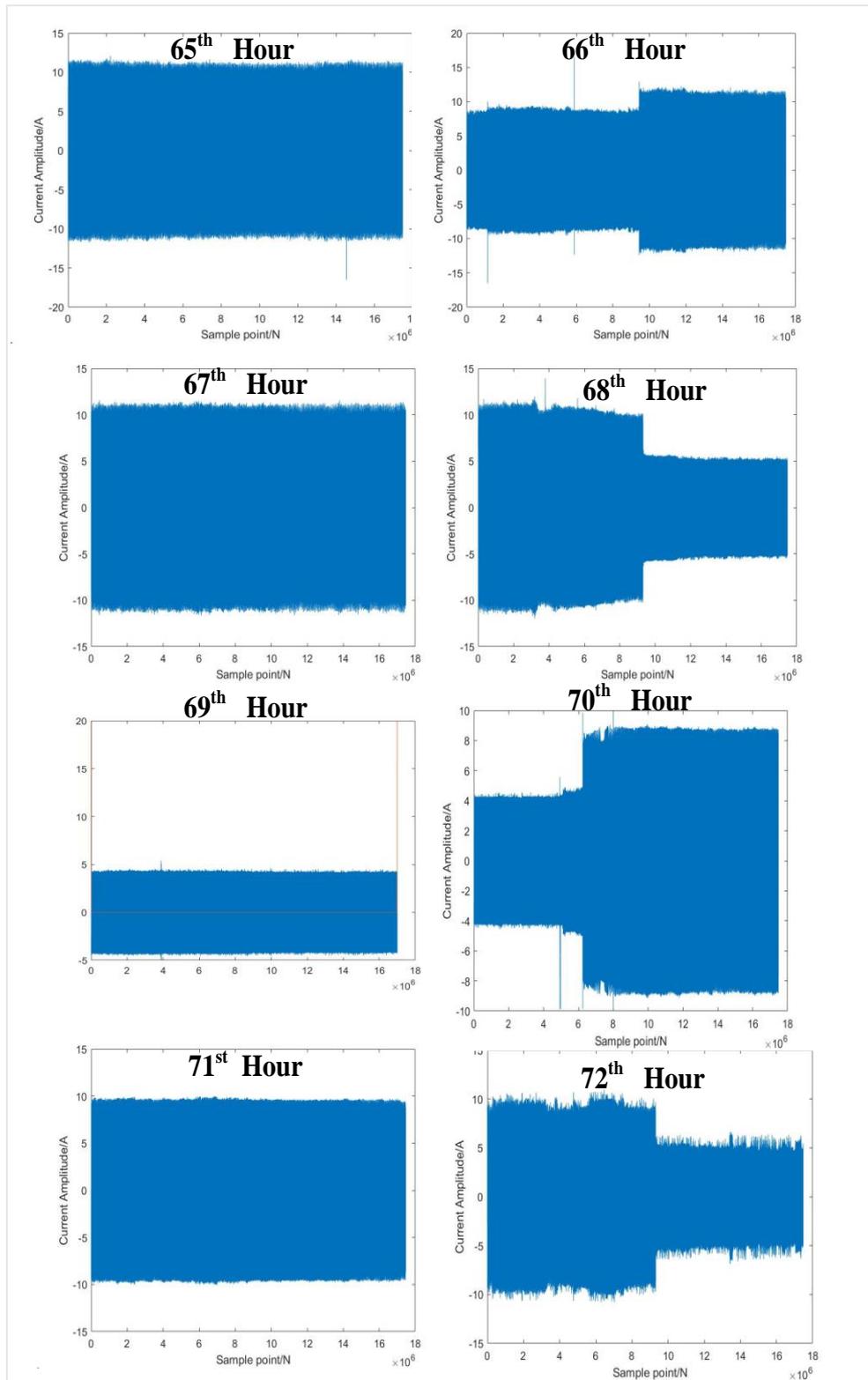
**Total current of office from 41<sup>st</sup> hour to 48<sup>th</sup> hour**



**Total current of office from 49<sup>th</sup> hour to 56<sup>th</sup> hour**

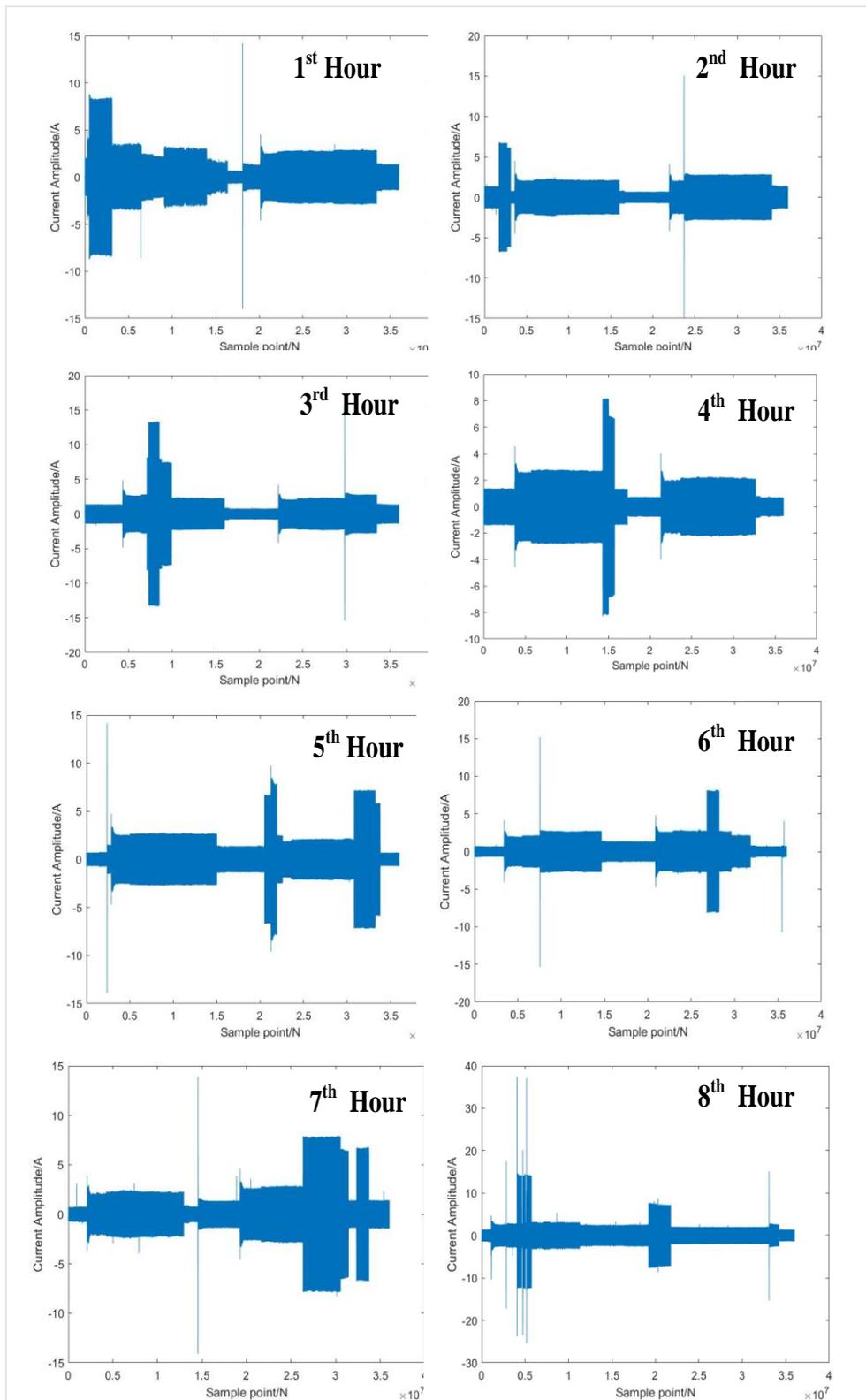


**Total current of office from 57<sup>th</sup> hour to 64<sup>th</sup> hour**

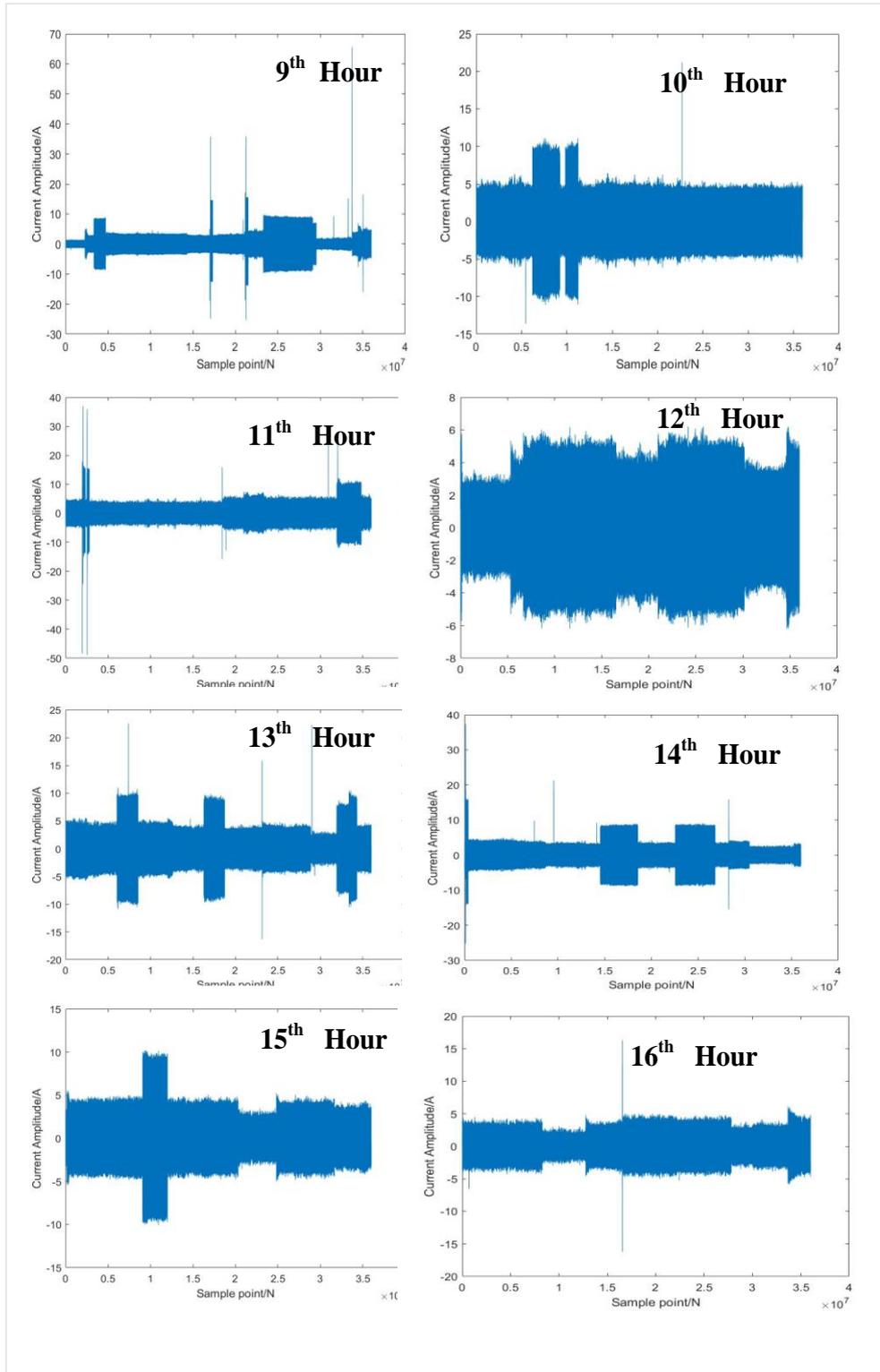


**Total current of office from 65<sup>th</sup> hour to 72<sup>th</sup> hour**

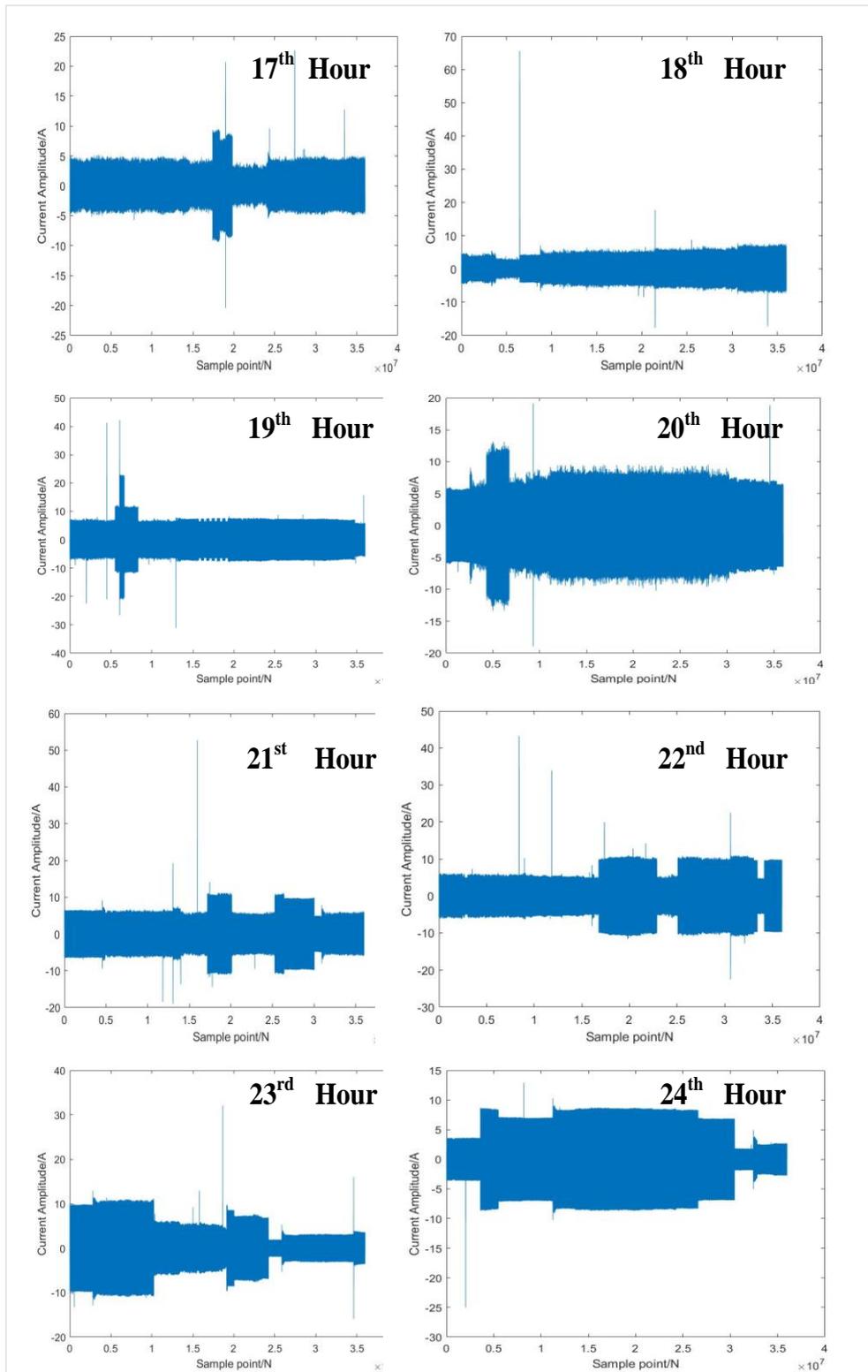
## Appendix II Data: Home B, Home C and Home D current waveform



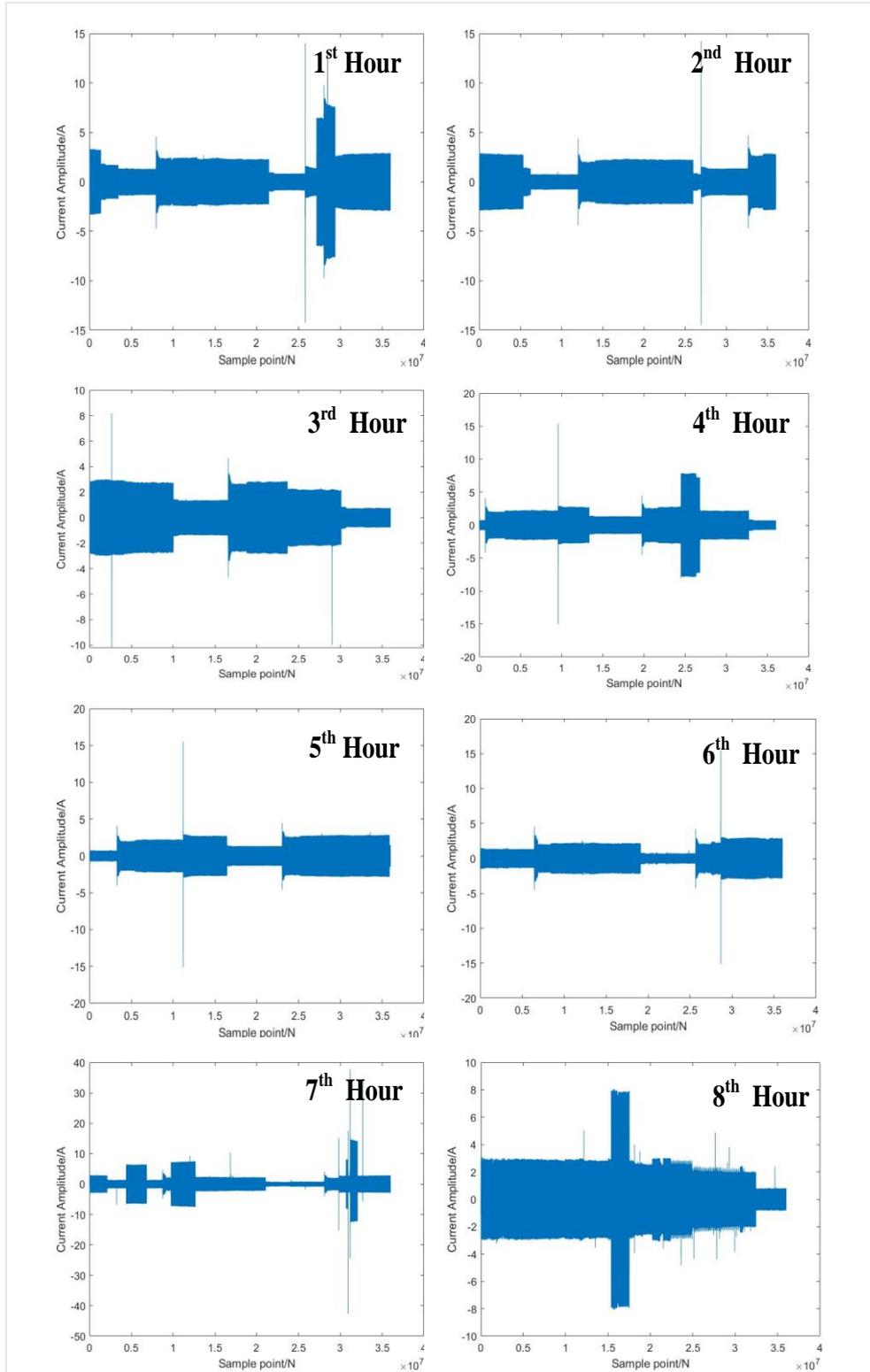
Total current of Home B from 1<sup>st</sup> hour to 8<sup>th</sup> hour



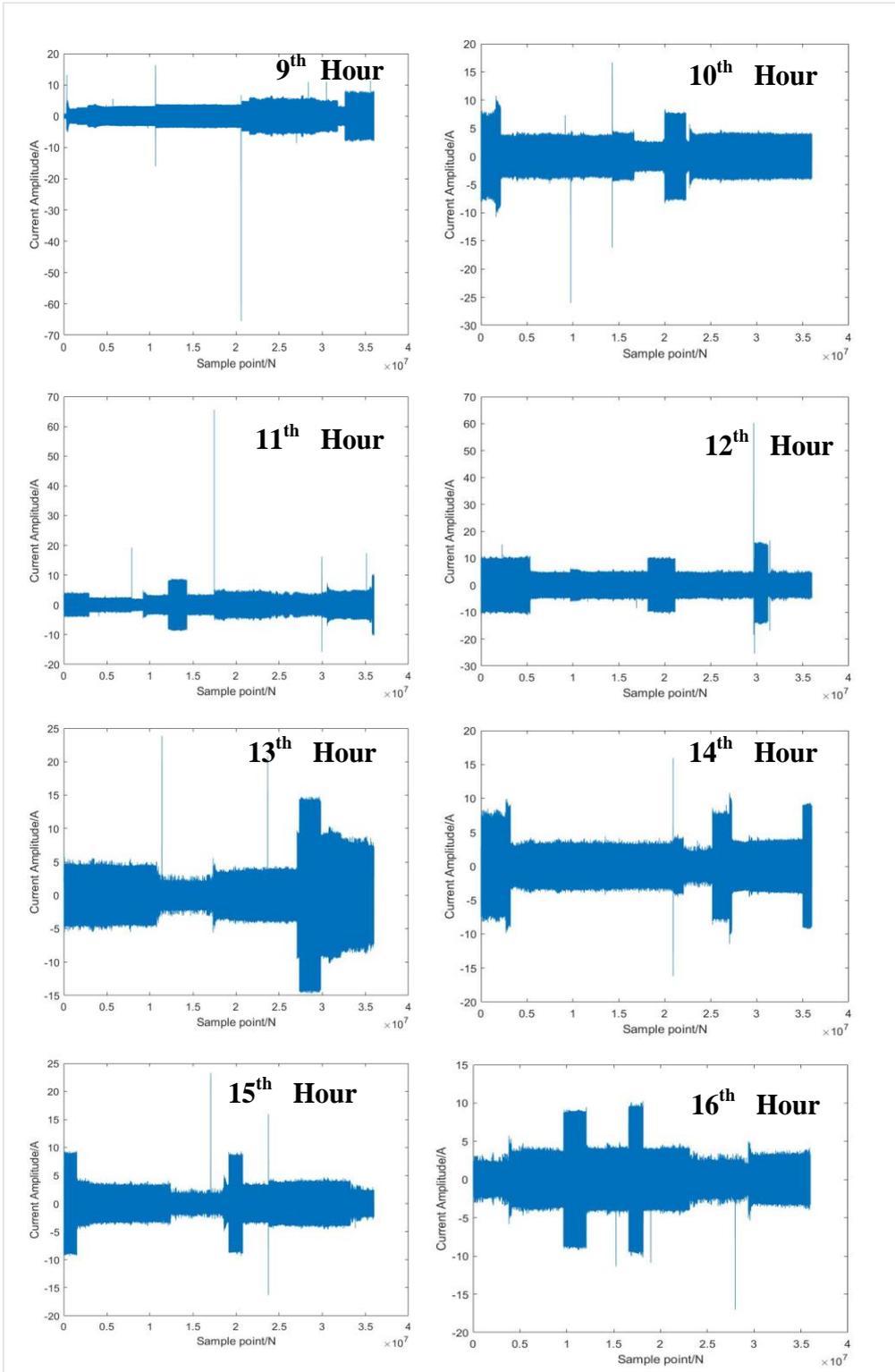
**Total current of Home B from 9<sup>th</sup> hour to 16<sup>th</sup> hour**



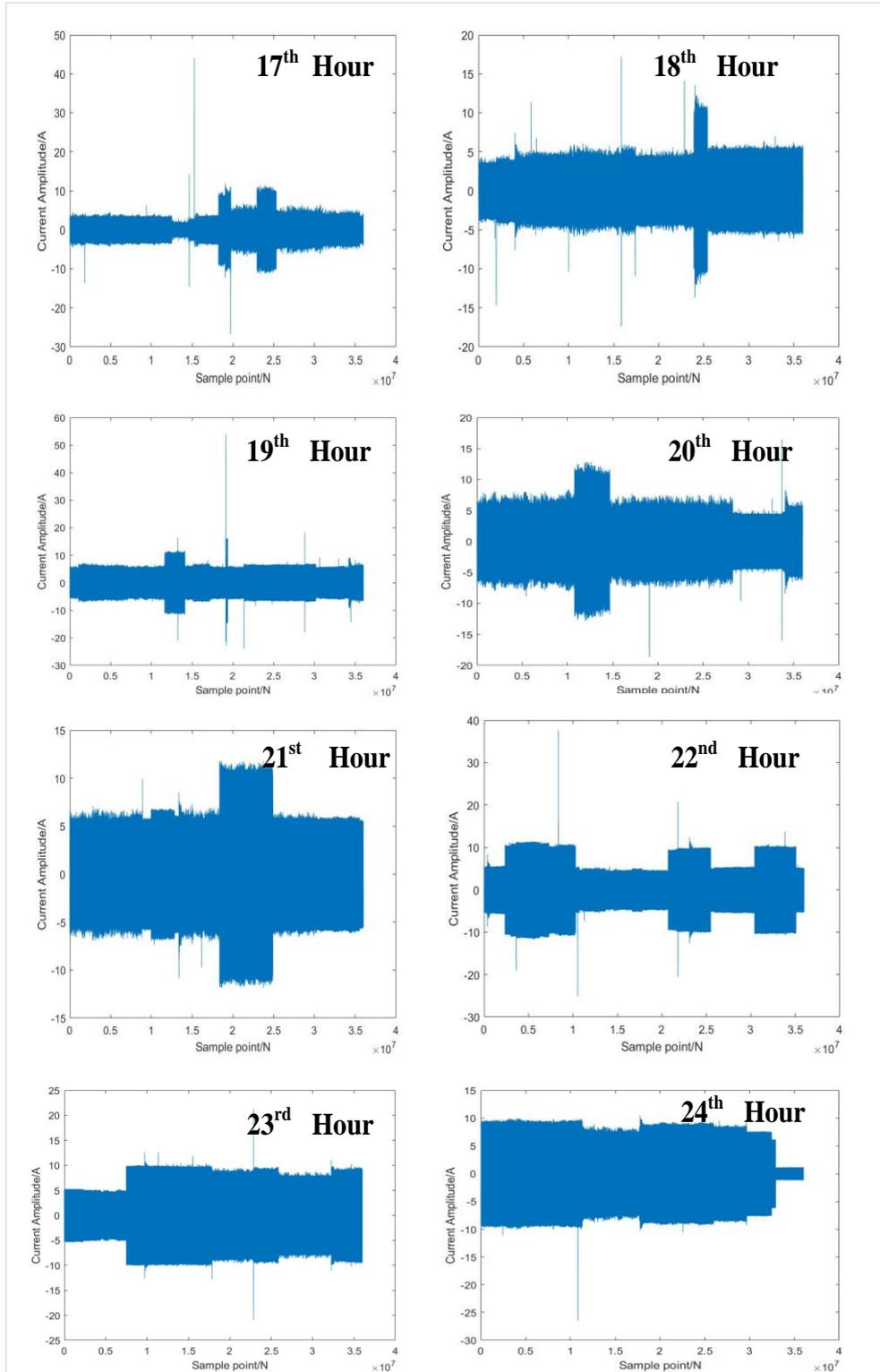
**Total current of Home B from 17<sup>th</sup> hour to 24<sup>th</sup> hour**



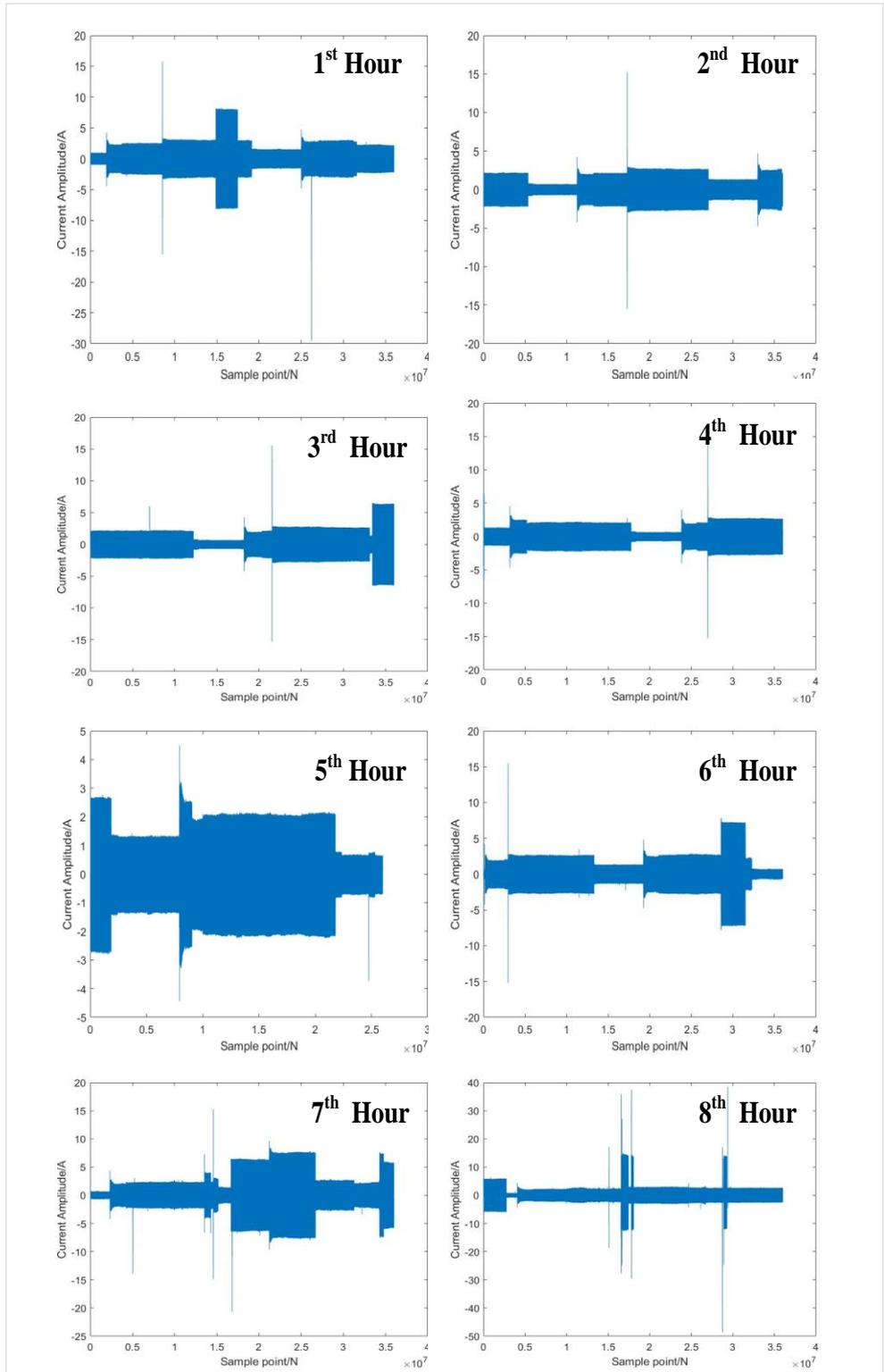
**Total current of Home C from 1<sup>st</sup> hour to 8<sup>th</sup> hour**



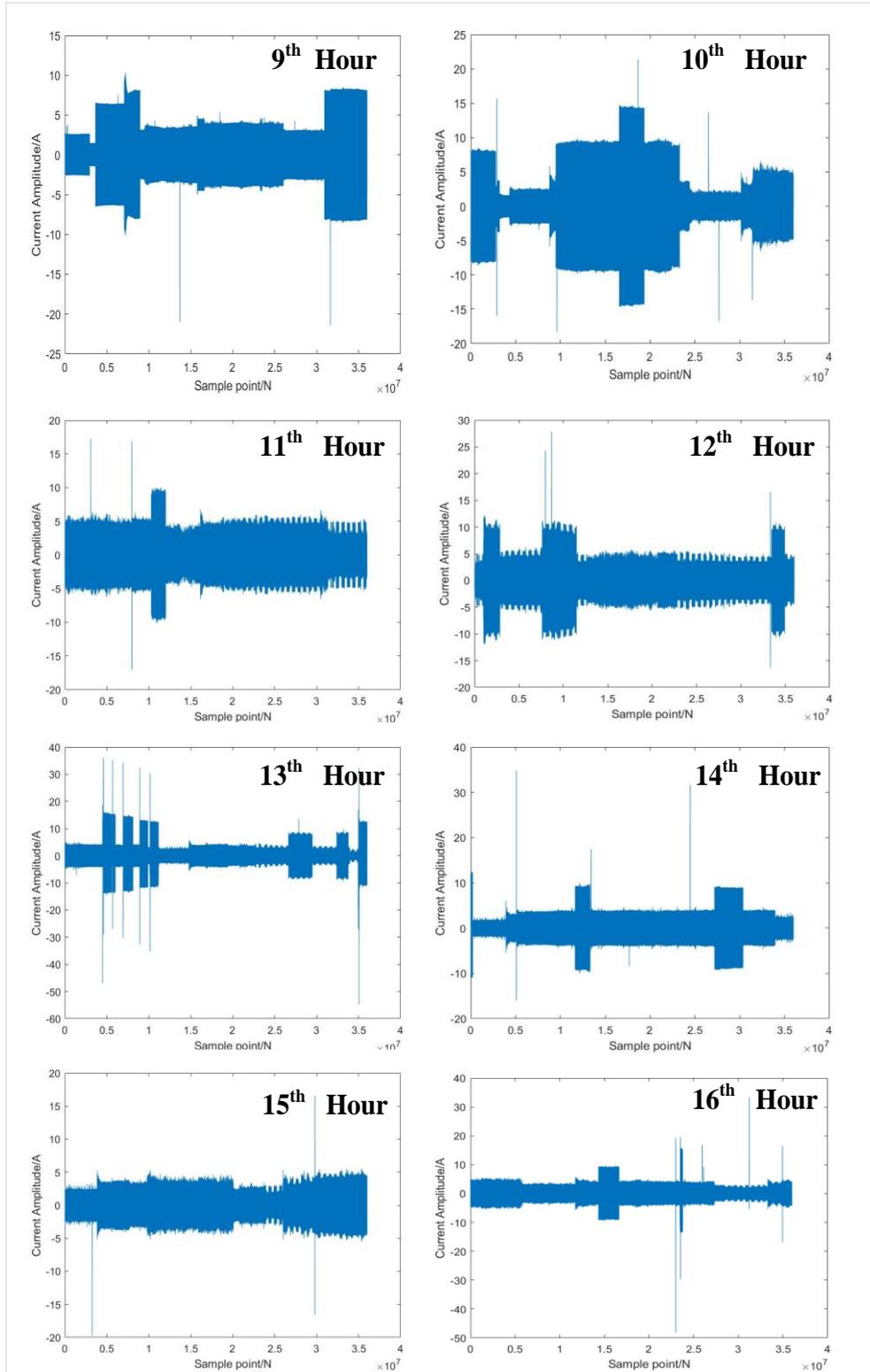
**Total current of Home C from 9<sup>th</sup> hour to 16<sup>th</sup> hour**



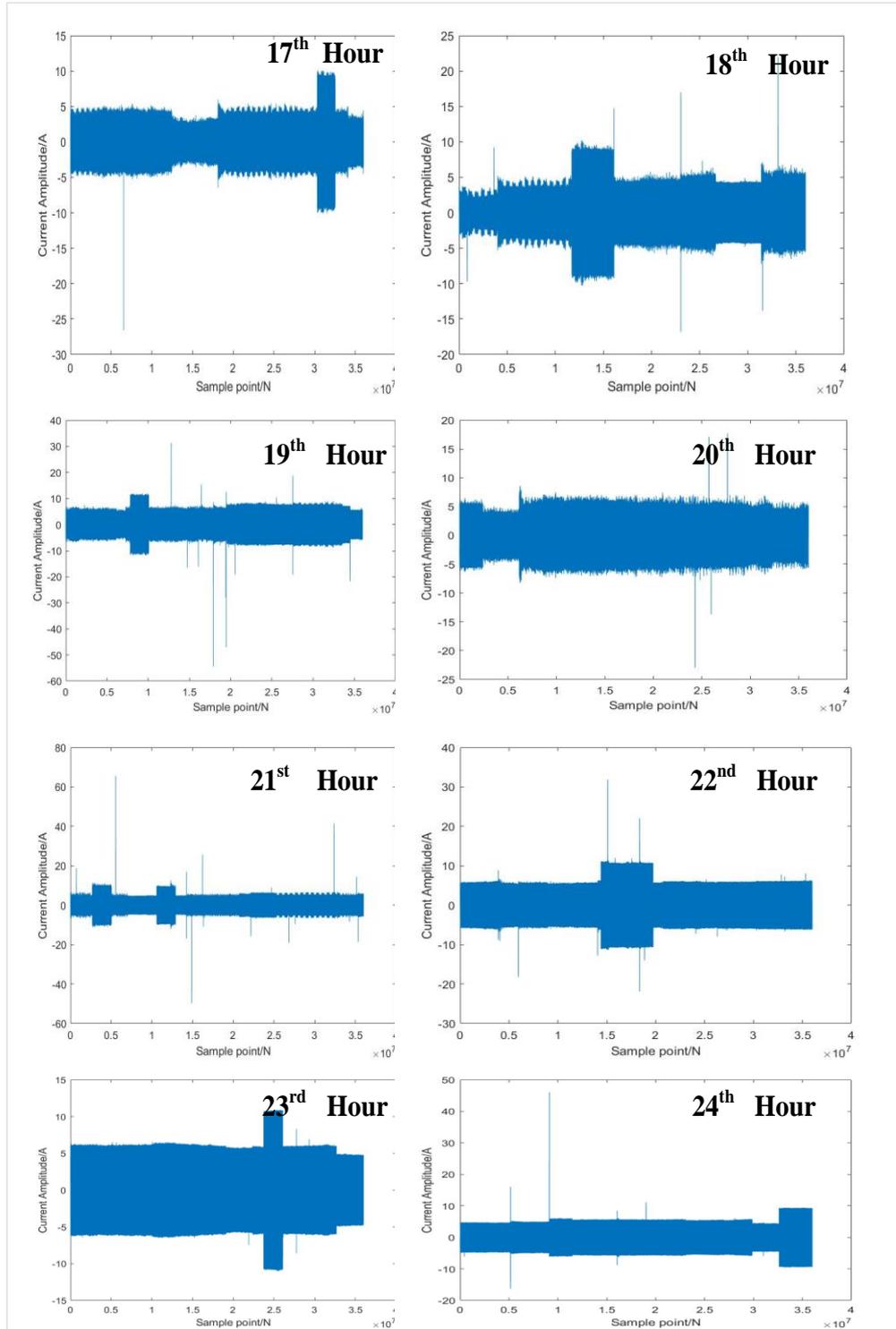
**Total current of Home C from 17<sup>th</sup> hour to 24<sup>th</sup> hour**



**Total current of Home D from 1<sup>st</sup> hour to 8<sup>th</sup> hour**



**Total current of Home D from 9<sup>th</sup> hour to 16<sup>th</sup> hour**



**Total current of Home D from 17<sup>th</sup> hour to 24<sup>th</sup> hour**

**Appendix III: The identification result of appliance switching in Home B, Home C and and Home D**

**Table 1 The identification result of appliance switching in Home B**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
1	00:05:23	TV switch off	62	11:51:12	Microwave oven switch off
2	00:08:45	Laptop switch off	63	12:24:15	Refrigerator switch on
3	00:13:56	Refrigerator switch on	64	12:32:11	Refrigerator switch off
4	00:41:35	Refrigerator switch off	65	12:57:39	Geyser switch on
5	00:42:23	Unknown	66	12:58:50	Geyser switch off
6	00:43:12	Unknown	67	13:12:45	Refrigerator switch on
7	00:47:23	Geyser switch on	68	13:17:18	Refrigerator switch off
8	01:04:34	Geyser switch off	69	13:55:02	TV switch off
9	01:05:38	Refrigerator switch on	70	14:02:00	Air-conditioning switch off
10	01:08:45	Geyser switch on	71	14:16:22	Refrigerator switch on
11	01:10:22	Noise	72	14:21:41	Refrigerator switch off
12	01:11:41	Geyser switch off	73	14:40:37	Air-conditioning switch on
13	01:12:56	Refrigerator switch off	74	14:51:41	Laptop switch on
13	01:12:56	Refrigerator switch off	74	14:51:41	Laptop switch on
14	01:54:13	Refrigerator switch on	75	15:03:20	Refrigerator switch on
15	01:58:35	Refrigerator switch off	76	15:09:59	TV switch off
16	02:03:35	Unknown	77	15:18:34	Refrigerator switch off
17	02:22:52	Refrigerator switch on	78	15:34:55	Laptop switch off
18	02:32:41	Refrigerator switch off	79	15:47:50	Air-conditioning switch off

**Table 1 The identification result of appliance switching in Home B**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
19	03:19:11	Refrigerator switch on	80	15:56:24	Refrigerator switch on
20	03:27:24	Refrigerator switch off	81	15:59:43	Refrigerator switch off
21	03:44:46	Geyser switch on	82	16:54:10	TV switch on
22	03:47:58	Refrigerator switch on	83	17:13:30	Kettle switch on
23	03:50:21	Geyser switch off	84	17:15:25	Kettle switch off
24	04:11:31	Refrigerator switch off	85	17:18:31	Refrigerator switch on
25	04:44:24	Refrigerator switch on	86	17:25:38	Refrigerator switch off
26	04:57:59	Refrigerator switch off	87	17:37:51	Laptop switch off
27	05:15:24	Refrigerator switch on	88	17:46:53	Laptop switch on
28	05:21:27	Refrigerator switch off	89	18:13:21	Laptop switch off
29	05:22:25	Noise	90	18:14:42	Refrigerator switch on
30	05:28:18	Geyser switch on	91	18:20:33	Refrigerator switch off
31	05:32:26	Geyser switch off	92	18:25:20	Geyser switch on
32	06:06:15	Refrigerator switch on	93	18:33:29	Geyser switch off
33	06:11:57	Refrigerator switch off	94	18:34:53	Air-conditioning switch on
34	07:28:45	Refrigerator switch on	95	18:34:46	Air-conditioning switch on
35	07:33:26	Refrigerator switch off	96	18:42:31	Unknown
36	08:27:32	Geyser switch on	97	18:49:37	TV switch off
37	08:33:37	Geyser switch off	98	19:10:39	Refrigerator switch on
38	08:45:47	Refrigerator switch on	99	19:20:51	Refrigerator switch off

**Table 1 The identification result of appliance switching in Home B**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
39	08:47:55	Geyser switch on	100	19:44:23	Microwave oven switch on
40	08:53:42	Refrigerator switch off	101	19:46:57	Kettle switch on
41	09:13:36	Geyser switch off	102	19:48:23	Kettle switch off
42	09:14:26	Unknown	103	19:50:21	Microwave oven switch off
43	09:19:51	Refrigerator switch on	104	20:27:25	Refrigerator switch on
44	09:22:36	Refrigerator switch off	105	20:35:37	Refrigerator switch off
45	09:37:41	Unknown	106	20:55:47	Geyser switch on
46	09:58:55	Refrigerator switch on	107	20:59:50	Geyser switch off
47	10:06:41	Refrigerator switch off	108	21:10:45	Refrigerator switch on
48	10:19:31	Microwave oven switch on	109	21:21:36	Refrigerator switch off
49	10:23:25	Microwave oven switch off	110	21:55:49	TV switch on
50	10:28:36	Geyser switch on	111	22:05:51	Air-conditioning switch off
51	10:32:27	Refrigerator switch on	112	22:13:03	Refrigerator switch on
52	10:36:48	Geyser switch off	113	22:21:41	Refrigerator switch off
53	10:37:51	Refrigerator switch off	114	22:48:20	Air-conditioning switch on
54	10:43:42	Air-conditioning switch on	115	22:50:41	Laptop switch on
55	10:49:39	TV switch on	116	23:01:21	Refrigerator switch on
56	11:10:29	Kettle switch on	117	23:06:00	Refrigerator switch off
57	11:12:24	Kettle switch off	118	23:09:51	TV switch off
58	11:24:33	Refrigerator switch on	119	23:35:36	Refrigerator switch on
59	11:33:56	Refrigerator switch off	120	23:47:34	Refrigerator switch off
60	11:44:27	Microwave oven switch on	121	23:50:41	Air-conditioning switch off
61	11:46:31	Kettle switch on			

**Table 2 The identification result of appliance switching in Home C**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
1	12:23:51	Refrigerator switch on	53	23:35:36	Refrigerator switch on
2	12:35:31	Refrigerator switch off	54	23:47:34	Refrigerator switch off
3	12:55:37	Geyser switch on	55	23:50:41	Unknown
4	12:57:05	Geyser switch off	56	01:30:46	Refrigerator switch on
5	13:20:54	Refrigerator switch on	57	01:35:23	Refrigerator switch off
6	13:23:48	Refrigerator switch off	58	02:41:46	Refrigerator switch on
7	13:50:22	TV switch on	59	02:45:23	Refrigerator switch off
8	14:01:56	Refrigerator switch on	60	03:18:03	Refrigerator switch on
9	14:17:00	Refrigerator switch off	61	03:29:16	Refrigerator switch off
10	14:20:14	Microwave oven switch on	62	03:43:55	Geyser switch on
11	14:43:26	Microwave oven switch off	63	03:46:38	Refrigerator switch on
12	14:55:14	TV switch off	64	03:49:05	Geyser switch off
13	15:01:02	Kettle switch on	65	04:10:03	Refrigerator switch off
14	15:05:40	Kettle switch off	66	05:17:41	Refrigerator switch on
15	15:08:58	Laptop switch off	67	05:23:28	Refrigerator switch off
16	15:16:30	Air-conditioning switch on	68	05:24:51	Noise
17	15:19:34	Refrigerator switch on	69	05:29:29	Geyser switch on
18	15:38:25	Refrigerator switch off	70	05:33:36	Geyser switch off
19	15:47:50	Noise	71	06:10:51	Refrigerator switch on
20	15:51:24	Geyser switch on	72	06:15:57	Refrigerator switch off
21	15:55:40	Geyser switch off	73	06:57:16	Refrigerator switch on
22	16:37:46	Kettle switch on	74	07:02:53	Refrigerator switch off

**Table 2 The identification result of appliance switching in Home C**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
23	16:38:32	Refrigerator switch on	75	07:21:50	Unknown
24	16:40:14	Kettle switch off	76	07:34:21	Unknown
25	16:45:04	Refrigerator switch off	77	07:44:10	Geyser switch on
26	17:17:02	TV switch on	78	08:10:19	Refrigerator switch on
27	17:20:36	Air-conditioning switch on	79	08:15:31	Refrigerator switch off
28	17:35:52	Refrigerator switch on	80	08:20:35	Microwave oven switch on
29	17:46:48	Refrigerator switch off	81	08:25:42	Microwave oven switch off
30	18:41:12	Microwave oven switch on	82	08:29:47	Laptop switch on
31	18:44:05	Microwave oven switch off	83	08:35:30	Unknown
32	18:18:00	Refrigerator switch on	84	08:45:37	Unknown
33	18:20:56	Refrigerator switch off	85	08:54:45	Refrigerator switch on
34	19:23:58	Noise	86	08:59:13	Refrigerator switch off
35	19:31:26	Geyser switch on	87	09:21:30	Microwave oven switch on
36	19:43:22	Geyser switch off	88	09:22:25	Microwave oven switch off
37	19:45:28	Air-conditioning switch off	89	09:27:28	Refrigerator switch on
38	20:27:25	Refrigerator switch on	90	09:30:38	Refrigerator switch off
39	20:35:37	Refrigerator switch off	91	09:45:26	Laptop switch on
40	20:55:47	Kettle switch on	92	10:08:16	Unknown
41	20:59:50	Kettle switch off	93	10:28:58	Refrigerator switch on
42	21:10:45	Refrigerator switch on	94	10:34:22	Refrigerator switch off
43	21:21:36	Refrigerator switch off	95	10:37:02	Geyser switch on

**Table 2 The identification result of appliance switching in Home C**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
44	21:55:49	TV switch on	96	10:41:26	Geyser switch off
45	22:05:51	Air-conditioning switch off	97	10:45:56	Air-conditioning switch on
46	22:13:03	Refrigerator switch on	98	10:47:48	TV switch on
47	22:21:41	Refrigerator switch off	99	11:07:38	Refrigerator switch on
48	22:48:20	Air-conditioning switch on	99	11:10:08	Refrigerator switch off
49	22:50:41	Laptop switch on	100	12:26:44	Refrigerator switch on
50	23:01:21	Refrigerator switch on	101	12:30:48	Refrigerator switch off
51	23:06:00	Refrigerator switch off	102	12:43:06	TV switch off
52	23:09:51	TV switch off	103	12:55:15	Laptop switch off

**Table 3 The identification result of appliance switching in Home D**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
1	20:11:32	TV switch on	55	09:15:53	Refrigerator switch off
2	20:15:46	Air-conditioning switch on	56	09:24:15	Refrigerator switch on
3	20:27:52	Refrigerator switch on	57	09:25:36	Air-conditioning switch on
4	20:36:18	Refrigerator switch off	58	09:34:14	Unknown
5	20:39:17	Laptop switch on	59	09:58:15	Laptop switch on
6	20:45:41	Unknown	60	10:07:21	Unknown
7	20:45:47	Unknown	61	10:18:43	Refrigerator switch on
8	21:31:21	Refrigerator switch on	62	10:23:25	Refrigerator switch off
9	21:43:50	Refrigerator switch off	63	10:30:36	Geyser switch on
10	22:28:00	Refrigerator switch on	64	10:33:16	Kettle switch on

**Table 3 The identification result of appliance switching in Home D**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
11	22:36:13	Refrigerator switch off	65	10:38:27	Kettle switch off
12	22:40:38	TV switch off	66	10:41:20	Geyser switch off
13	22:43:37	Geyser switch on	67	10:44:42	Unknown
14	22:45:24	Geyser switch off	68	10:51:48	TV switch on
15	22:50:30	Air-conditioning switch off	69	11:13:39	Kettle switch on
16	22:55:34	Refrigerator switch on	70	11:15:42	Kettle switch off
17	23:02:15	Refrigerator switch off	71	11:34:33	Refrigerator switch on
18	23:32:42	Unknown	72	11:43:15	Refrigerator switch off
19	00:40:22	Refrigerator switch on	73	11:46:38	Microwave oven switch on
20	00:47:01	Refrigerator switch off	74	11:49:13	Microwave oven switch off
21	00:43:45	Unknown	75	12:35:24	Refrigerator switch on
22	00:47:16	Geyser switch on	76	12:39:27	Refrigerator switch off
23	01:13:40	Geyser switch off	77	12:53:59	TV switch off
24	01:15:29	Refrigerator switch on	78	12:59:51	Laptop switch off
25	01:19:32	Geyser switch on	79	13:01:56	Air-conditioning switch on
26	01:20:40	Noise	80	13:07:00	Unknown
27	01:22:21	Geyser switch off	81	13:21:43	Refrigerator switch on
28	01:26:20	Refrigerator switch off	82	13:33:15	Refrigerator switch off
29	01:58:42	Unknown	83	13:55:41	TV switch on
30	02:04:11	Unknown	84	14:21:23	Refrigerator switch on
31	02:07:45	Unknown	85	14:35:04	Refrigerator switch off
32	02:47:31	Refrigerator switch on	86	15:06:18	TV switch off

**Table 3 The identification result of appliance switching in Home D**

<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>	<b>NO.</b>	<b>Time Point</b>	<b>Identified Result</b>
33	02:50:37	Refrigerator switch off	87	15:07:38	Laptop switch on
34	03:27:13	Refrigerator switch on	88	16:25:47	Refrigerator switch on
35	03:39:25	Refrigerator switch off	89	16:35:32	Refrigerator switch off
36	03:53:15	Geyser switch on	90	16:55:47	Air-conditioning switch on
37	03:56:47	Refrigerator switch on	91	16:58:03	Kettle switch on
37	03:56:47	Refrigerator switch on	91	16:58:03	Kettle switch on
38	04:04:15	Geyser switch off	92	17:01:16	Kettle switch on
39	04:05:37	Refrigerator switch off	93	17:19:13	Refrigerator switch on
50	08:23:52	Geyser switch off	103	19:53:35	Laptop switch off
51	08:34:36	Refrigerator switch on	104	20:24:52	Refrigerator switch on
52	08:35:41	Geyser switch on	105	20:36:26	Refrigerator switch off
53	08:42:21	Refrigerator switch off	106	20:45:39	Kettle switch on
54	09:03:52	Geyser switch off	107	20:48:05	Kettle switch off

## Reference

1. *World electricity consumption* Net electricity consumption worldwide in select years, <https://www.statista.com/statistics/280704/world-power-consumption/>.
2. Goyal, G.R. and S. Vadhera, *Challenges of implementing demand side management in developing countries*. Journal of Power Technologies, 2020. **100**(1): p. 43.
3. Schirmer, P.A. and I. Mporas, *Non-Intrusive Load Monitoring: A Review*. IEEE Transactions on Smart Grid, 2022.
4. Mahdavi, N. and J.H. Braslavsky, *Modelling and control of ensembles of variable-speed air conditioning loads for demand response*. IEEE Transactions on Smart Grid, 2020. **11**(5): p. 4249-4260.
5. Shah, Z.A., et al., *Fuzzy logic-based direct load control scheme for air conditioning load to reduce energy consumption*. IEEE Access, 2020. **8**: p. 117413-117427.
6. Ding, Y., et al., *Game-theoretic demand side management of thermostatically controlled loads for smoothing tie-line power of microgrids*. IEEE Transactions on Power Systems, 2021. **36**(5): p. 4089-4101.
7. Jiang, T., et al., *Coordinated control of air-conditioning loads for system frequency regulation*. IEEE Transactions on Smart Grid, 2020. **12**(1): p. 548-560.
8. Peirelinck, T., et al., *Domain randomization for demand response of an electric water heater*. IEEE Transactions on Smart Grid, 2020. **12**(2): p. 1370-1379.
9. Fu, L., et al., *Optimal restoration of an unbalanced distribution system into multiple microgrids considering three-phase demand-side management*. IEEE Transactions on Power Systems, 2020. **36**(2): p. 1350-1361.
10. Zhou, K. and S. Yang, *Demand side management in China: The context of China's power industry reform*. Renewable and Sustainable Energy Reviews, 2015. **47**: p. 954-965.
11. Hui, H., et al., *Operating reserve evaluation of aggregated air conditioners*. Applied energy, 2017. **196**: p. 218-228.
12. Gong, X., et al., *Robust hierarchical control mechanism for aggregated thermostatically controlled loads*. IEEE Transactions on Smart Grid, 2020. **12**(1): p. 453-467.
13. Gong, F., et al. *A svm optimized by particle swarm optimization approach to load disaggregation in non-intrusive load monitoring in smart homes*. in *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*. 2019. IEEE.
14. Yang, D., et al., *An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance*. IEEE Transactions on Consumer Electronics, 2020. **66**(2): p. 173-182.
15. Du, Z. and G. Yan. *Two-Step Non-Intrusive Load Monitoring Based on High-Frequency Current Data*. in *2020 39th Chinese Control Conference (CCC)*. 2020. IEEE.
16. Kim, J.-G. and B. Lee, *Appliance classification by power signal analysis based on multi-feature combination multi-layer LSTM*. Energies, 2019. **12**(14): p. 2804.
17. Figueiredo, M.B., A.d. Almeida, and B. Ribeiro. *An experimental study on electrical signature identification of non-intrusive load monitoring (nilm) systems*. in *International Conference on Adaptive and Natural Computing Algorithms*. 2011. Springer.
18. Ma, C. and L. Yin, *Deep Flexible Transmitter Networks for Non-Intrusive Load Monitoring of Power Distribution Networks*. IEEE Access, 2021. **9**: p. 107424-107436.
19. Ghosh, S. and D. Chatterjee, *Artificial bee colony optimization based non-intrusive appliances load monitoring technique in a smart home*. IEEE Transactions on Consumer Electronics, 2021. **67**(1): p. 77-86.
20. Drenker, S. and A. Kader, *Nonintrusive monitoring of electric loads*. IEEE Computer Applications in Power, 1999. **12**(4): p. 47-51.
21. Kelly, J. and W. Knottenbelt, *The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes*. Scientific data, 2015. **2**(1): p. 1-14.
22. Buddhahai, B., W. Wongseeree, and P. Rakkwamsuk, *An energy prediction approach for a nonintrusive load monitoring in home appliances*. IEEE Transactions on Consumer Electronics,

2019. **66**(1): p. 96-105.
23. Jaramillo, A.F.M., et al., *Load modelling and non-intrusive load monitoring to integrate distributed energy resources in low and medium voltage networks*. Renewable Energy, 2021. **179**: p. 445-466.
  24. Hart, G.W., *Nonintrusive appliance load monitoring*. Proceedings of the IEEE, 1992. **80**(12): p. 1870-1891.
  25. Himeur, Y., et al., *Recent trends of smart nonintrusive load monitoring in buildings: A review, open challenges, and future directions*. International Journal of Intelligent Systems, 2022. **37**(10): p. 7124-7179.
  26. Lu, J., et al., *An Overview of Non-Intrusive Load Monitoring Based on VI Trajectory Signature*. Energies, 2023. **16**(2): p. 939.
  27. Guo, L., et al., *A load identification method based on active deep learning and discrete wavelet transform*. IEEE Access, 2020. **8**: p. 113932-113942.
  28. Nie, Z., Y. Yang, and Q. Xu, *An ensemble-policy non-intrusive load monitoring technique based entirely on deep feature-guided attention mechanism*. Energy and Buildings, 2022. **273**: p. 112356.
  29. Mukaroh, A., T.-T.-H. Le, and H. Kim, *Background load denoising across complex load based on generative adversarial network to enhance load identification*. Sensors, 2020. **20**(19): p. 5674.
  30. Piccialli, V. and A.M. Sudoso, *Improving non-intrusive load disaggregation through an attention-based deep neural network*. Energies, 2021. **14**(4): p. 847.
  31. Hadi, M.U., N.H.N. Suhaimi, and A. Basit, *Efficient Supervised Machine Learning Network for Non-Intrusive Load Monitoring*. Technologies, 2022. **10**(4): p. 85.
  32. Chekan, J.A. and S. Bashash. *IOT-oriented control of aggregate thermostatically-controlled loads*. in *2017 American Control Conference (ACC)*. 2017. IEEE.
  33. Gibbs, E., *Foreign direct investment policy, multinationals, and subsidiary entrepreneurship success and failure in post-war Scotland*. Business History, 2022: p. 1-23.
  34. Malmedal, K., B. Kroposki, and P.K. Sen. *Energy Policy Act of 2005 and its impact on renewable energy applications in USA*. in *2007 IEEE Power Engineering Society General Meeting*. 2007. IEEE.
  35. Waxman, H.A. and E.J. Markey, *American clean energy and security act of 2009*. Washington: US House of Representatives, 2009.
  36. Pollitt, M.G., *Lessons from the history of independent system operators in the energy sector*. Energy Policy, 2012. **47**: p. 32-48.
  37. Greenfield, L.R., *An Overview of the Federal Energy Regulatory Commission and Federal Regulation of Public Utilities*. 2018.
  38. Khanna, N.Z., J. Guo, and X. Zheng, *Effects of demand side management on Chinese household electricity consumption: Empirical findings from Chinese household survey*. Energy Policy, 2016. **95**: p. 113-125.
  39. 戴庆忠, *优化电源结构走电力可持续发展道路——"十二五"电力工业发展规划解析*. 东方电机, 2013. **000**(002): p. 1-16.  
Dai Qingzhong, *optimize the power structure and take the path of sustainable development of electric power -- Analysis of the 12th Five Year Plan for the development of electric power industry* Dongfang Electric machinery, 2013. **000** (002): P. 1-16.
  40. 无, *新版《电力并网运行管理规定》《电力辅助服务管理办法》发布*. 电力设备管理, 2022(1): p. 3.  
None, the new version of "Regulations on the administration of grid connected operation of electric power" and "measures for the administration of auxiliary services of electric power" were issued Power equipment management, 2022 (1): P. 3.
  41. Wu, Y.-K., et al., *The effect of decision analysis on power system resilience and economic value during a severe weather event*. IEEE Transactions on Industry Applications, 2022. **58**(2): p. 1685-1695.
  42. Azizi, E., et al., *Residential household non-intrusive load monitoring via smart event-based optimization*. IEEE Transactions on Consumer Electronics, 2020. **66**(3): p. 233-241.
  43. Heo, S. and H. Kim, *Toward load identification based on the hilbert transform and sequence*

- to sequence long short-term memory. *IEEE Transactions on Smart Grid*, 2021. **12**(4): p. 3252-3264.
44. Chang, H.-H., et al., *Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification*. *IEEE Transactions on Industry Applications*, 2013. **50**(3): p. 2081-2089.
  45. Klemenjak, C., D. Egarter, and W. Elmenreich, *YoMo: the Arduino-based smart metering board*. *Computer Science-Research and Development*, 2016. **31**(1): p. 97-103.
  46. Makonin, S., et al. *Inspiring energy conservation through open source metering hardware and embedded real-time load disaggregation*. in *2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. 2013. IEEE.
  47. Azizi, E., M.T. Beheshti, and S. Bolouki, *Event matching classification method for non-intrusive load monitoring*. *Sustainability*, 2021. **13**(2): p. 693.
  48. Khazaei, M., L. Stankovic, and V. Stankovic. *Evaluation of low-complexity supervised and unsupervised NILM methods and pre-processing for detection of multistate white goods*. in *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*. 2020.
  49. Fang, K., et al. *An event detection approach based on improved CUSUM algorithm and kalman filter*. in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*. 2020. IEEE.
  50. Houidi, S., et al., *Multivariate event detection methods for non-intrusive load monitoring in smart homes and residential buildings*. *Energy and Buildings*, 2020. **208**: p. 109624.
  51. Pereira, L. *Developing and evaluating a probabilistic event detector for non-intrusive load monitoring*. in *2017 Sustainable Internet and ICT for Sustainability (SustainIT)*. 2017. IEEE.
  52. Lu, M. and Z. Li, *A hybrid event detection approach for non-intrusive load monitoring*. *IEEE Transactions on Smart Grid*, 2019. **11**(1): p. 528-540.
  53. Pereira, L. and N. Nunes, *An empirical exploration of performance metrics for event detection algorithms in non-intrusive load monitoring*. *Sustainable Cities and Society*, 2020. **62**: p. 102399.
  54. Drouaz, M., et al., *New time-frequency transient features for nonintrusive load monitoring*. *Energies*, 2021. **14**(5): p. 1437.
  55. de Aguiar, E.L., et al., *Scattering transform for classification in non-intrusive load monitoring*. *Energies*, 2021. **14**(20): p. 6796.
  56. Liu, H., H. Wu, and C. Yu, *A hybrid model for appliance classification based on time series features*. *Energy and Buildings*, 2019. **196**: p. 112-123.
  57. Li, X., et al. *A training-free non-intrusive load monitoring approach for high-frequency measurements based on graph signal processing*. in *2022 7th Asia Conference on Power and Electrical Engineering (ACPEE)*. 2022. IEEE.
  58. Ortega, A., et al., *Graph signal processing: Overview, challenges, and applications*. *Proceedings of the IEEE*, 2018. **106**(5): p. 808-828.
  59. Himeur, Y., et al., *Smart non-intrusive appliance identification using a novel local power histogramming descriptor with an improved k-nearest neighbors classifier*. *Sustainable Cities and Society*, 2021. **67**: p. 102764.
  60. Himeur, Y., et al., *Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors*. *International Journal of Intelligent Systems*, 2021. **36**(6): p. 2865-2894.
  61. Hernandez, A.S., A.H. Ballado, and A.P.D. Heredia. *Development of a non-intrusive load monitoring (nilm) with unknown loads using support vector machine*. in *2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS)*. 2021. IEEE.
  62. Khan, M.M.R., M.A.B. Siddique, and S. Sakib. *Non-intrusive electrical appliances monitoring and classification using K-nearest neighbors*. in *2019 2nd International Conference on Innovation in Engineering and Technology (ICIET)*. 2019. IEEE.
  63. Verma, A., et al., *A comprehensive review on the nilm algorithms for energy disaggregation*. *arXiv preprint arXiv:2102.12578*, 2021.
  64. Dash, S. and N. Sahoo, *Electric energy disaggregation via non-intrusive load monitoring: A state-of-the-art systematic review*. *Electric Power Systems Research*, 2022. **213**: p. 108673.

65. Ya-Ting, H., et al., *Research advances and perspectives on the cocktail party problem and related auditory models*. Acta Automatica Sinica, 2019. **45**(2): p. 234-251.
66. Feng, F. and M. Kowalski, *Revisiting sparse ICA from a synthesis point of view: Blind Source Separation for over and underdetermined mixtures*. Signal Processing, 2018. **152**: p. 165-177.
67. Xiao, Z., et al., *Comparison between artificial neural network and random forest for effective disaggregation of building cooling load*. Case Studies in Thermal Engineering, 2021. **28**: p. 101589.
68. Liu, Y., et al., *Non-intrusive load monitoring based on unsupervised optimization enhanced neural network deep learning*. Frontiers in Energy Research, 2021. **9**: p. 718916.
69. Ciancetta, F., et al., *A new convolutional neural network-based system for NILM applications*. IEEE Transactions on Instrumentation and Measurement, 2020. **70**: p. 1-12.
70. Yang, L., J.a. Chen, and W. Zhu, *Dynamic hand gesture recognition based on a leap motion controller and two-layer bidirectional recurrent neural network*. Sensors, 2020. **20**(7): p. 2106.
71. Kalinke, F., et al. *An evaluation of nilm approaches on industrial energy-consumption data*. in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. 2021.
72. Qadeer, K., et al., *A long short-term memory (LSTM) network for hourly estimation of PM<sub>2.5</sub> concentration in two cities of South Korea*. Applied Sciences, 2020. **10**(11): p. 3984.
73. He, X., et al., *A Novel Denoising Auto-Encoder-Based Approach for Non-Intrusive Residential Load Monitoring*. Energies, 2022. **15**(6): p. 2290.
74. Garcia-Perez, D., et al., *Fully-convolutional denoising auto-encoders for NILM in large non-residential buildings*. IEEE Transactions on Smart Grid, 2020. **12**(3): p. 2722-2731.
75. Laouali, I.H., et al. *A survey on computational intelligence techniques for non intrusive load monitoring*. in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*. 2020. IEEE.
76. Iqbal, H.K., et al., *A critical review of state-of-the-art non-intrusive load monitoring datasets*. Electric Power Systems Research, 2021. **192**: p. 106921.
77. Harell, A., S. Makonin, and I.V. Bajić. *Wavenilm: A causal neural network for power disaggregation from the complex power signal*. in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019. IEEE.
78. Svensson, S., *Power Measurement Techniques for Nonsinusoidal Conditions: The Significance of Harmonics for the Measurement of Power and Other AC Quantities*. 1999: Chalmers Tekniska Hogskola (Sweden).
79. Bonfigli, R., et al., *Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models*. Applied Energy, 2017. **208**: p. 1590-1607.
80. Sadeghianpourhamami, N., et al., *Comprehensive feature selection for appliance classification in NILM*. Energy and Buildings, 2017. **151**: p. 98-106.
81. Dong, M., et al., *Non-intrusive signature extraction for major residential loads*. IEEE Transactions on Smart Grid, 2013. **4**(3): p. 1421-1430.
82. Ha, S.J., et al., *Panorama mosaic optimization for mobile camera systems*. IEEE transactions on consumer electronics, 2007. **53**(4): p. 1217-1225.
83. Gray, M. and W.G. Morsi. *Application of wavelet-based classification in non-intrusive load monitoring*. in *2015 IEEE 28th Canadian conference on electrical and computer engineering (CCECE)*. 2015. IEEE.
84. Tabatabaei, S.M., S. Dick, and W. Xu, *Toward non-intrusive load monitoring via multi-label classification*. IEEE Transactions on Smart Grid, 2016. **8**(1): p. 26-40.
85. Li, K., et al., *A Nonintrusive load identification model based on time-frequency features fusion*. IEEE Access, 2020. **9**: p. 1376-1387.
86. Puente, C., et al., *Non-intrusive load monitoring (NILM) for energy disaggregation using soft computing techniques*. Energies, 2020. **13**(12): p. 3117.
87. Lazzaretti, A.E., et al., *A multi-agent NILM architecture for event detection and load classification*. Energies, 2020. **13**(17): p. 4396.
88. Chan, W.L., A.T. So, and L. Lai. *Harmonics load signature recognition by wavelets transforms*. in *DRPT2000. International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. Proceedings (Cat. No. 00EX382)*. 2000. IEEE.
89. Su, Y.-C., K.-L. Lian, and H.-H. Chang. *Feature selection of non-intrusive load monitoring system*

- using STFT and wavelet transform. in *2011 IEEE 8th international conference on e-business engineering*. 2011. IEEE.
90. Miller, C., *More buildings make more generalizable Models—Benchmarking prediction methods on open electrical meter data*. Machine Learning and Knowledge Extraction, 2019. **1**(3): p. 974-993.
  91. Niesche, R. and M. Haase, *Emotions and ethics: A Foucauldian framework for becoming an ethical educator*. Educational philosophy and theory, 2012. **44**(3): p. 276-288.
  92. Ahmadi-Karvigh, S., et al., *Real-time activity recognition for energy efficiency in buildings*. Applied energy, 2018. **211**: p. 146-160.
  93. Gulati, M., S.S. Ram, and A. Singh, *Leveraging EMI signals for appliance detection and energy harvesting*. 2020, IIIT-Delhi.
  94. Kong, S., et al., *Home appliance load disaggregation using cepstrum-smoothing-based method*. IEEE Transactions on Consumer Electronics, 2015. **61**(1): p. 24-30.
  95. Meehan, P., C. McArdle, and S. Daniels, *An efficient, scalable time-frequency method for tracking energy usage of domestic appliances using a two-step classification algorithm*. Energies, 2014. **7**(11): p. 7041-7066.
  96. Gurbuz, F.B., R. Bayindir, and S. Vadi. *Comprehensive non-intrusive load monitoring process: Device event detection, device feature extraction and device identification using KNN, random forest and decision tree*. in *2021 10th International Conference on Renewable Energy Research and Application (ICRERA)*. 2021. IEEE.
  97. Feng, T., et al., *A Non-intrusive Load Monitoring Method Based on Improved kNN Algorithm and Transient Steady State Features*. Computer and Modernization, 2022(10): p. 29.
  98. Hajizadeh, R., A. Aghagolzadeh, and M. Ezoji, *Mutual neighborhood and modified majority voting based KNN classifier for multi-categories classification*. Pattern Analysis and Applications, 2022. **25**(4): p. 773-793.
  99. Yuan, D.H., B.S.H. Chew, and Y. Xu. *A Time Series Data Feature Based Back Propagation Neural Network Approach for Non-intrusive Load Monitoring*. in *2019 9th International Conference on Power and Energy Systems (ICPES)*. 2019. IEEE.
  100. Ye, K., et al., *Optimization of lapping process parameters of CP-Ti based on PSO with mutation and BPNN*. The International Journal of Advanced Manufacturing Technology, 2021. **117**(9-10): p. 2859-2866.
  101. Kleinberg, J., *The convergence of social and technological networks*. Communications of the ACM, 2008. **51**(11): p. 66-72.