DEFINING OF NORMALIZED LOAD PROFILE CURVES FOR DOMESTIC CUSTOMER GROUPS TO ESTIMATE FEEDER POWER LOSS

H.A.C.H.Jayawardhana

(139505 D)

Dissertation submitted in partial fulfillment of the requirements for the Degree Master of Science in Electrical Engineering

Department of Electrical Engineering

University of Moratuwa

Sri Lanka

May 2018

DECLARATION OF THE CANDIDATE AND SUPERVISOR

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

.... 08/05/2018

Signature of the candidate Date:

(H.A.C.H. Jayawardhana)

The above candidate has carried out research for the Masters Dissertation under my supervision.

UOM Verified Signature 08/05/2018

Signature of the supervisor **Date:**

(Dr. K.T.M. Udayanga Hemapala)

UOM Verified Signature 08/05/2018.

Signature of the supervisor **Date:**

(Dr. P.S. Narendra De Silva)

Abstract

'Estimation of load profiles for domestic customers' is a multi-purpose activity and 'Estimation of daily feeder power loss' is only a one use of customer load profiles. In a country, domestic electricity customer percentage is higher in number wise, but energy usage of one customer is lower compared to other categories. Therefore installation of load profile recording meters for each domestic customer to obtain customer load profile is impractical and not economical.

In this research, set of domestic customers are grouped by clustering their daily load profiles with respect to differences of patterns. Representative normalized load profile is obtained for each group. Same customers were interviewed for collecting family member composition and electric equipment usage information. Relationships between load profile pattern and customer information were investigated. Then a methodology was developed to estimate load profile of a new customer by only using customer information and monthly total energy consumption. These load profiles were used to calculate low voltage feeder power loss.

As outcome of this research, MATLAB GUI software interface was developed to input customer information and selection of best-matched representative load profile of a new customer. An algorithm is proposed to estimate time dependant LV feeder power loss by using estimated customer load profiles.

Acknowledgement

I would like to express my sincere gratitude to those who were behind me in completing this research project.

I am deeply indebted to my supervisors Dr. Narendra De Silva, Head of Engineering, Lanka Electricity Company and Dr. Udayanga Hemapala, Senior Lecturer, Department of Electrical Engineering, University of Moratuwa for their continuous support and encouragement from starting of this project up to writing this dissertation.

In addition, I would like to thank all the lecturers in Department of Electrical Engineering and Post Graduate Office, Faculty of Engineering, who engaged in this MSc Course in various ways to educate us and broaden our vision. Dr. Asoka Korale, Senior Technology Consultant, MIT helped me with his broad knowledge in algorithm development. Thank you very much for allocation time to educate me.

My sincere thanks go to my managers and colleagues at Lanka Electricity Company (Private) Limited who help me in many ways during this period. I specially thanking for colleagues in IT Department, Test Department and Kotte Branch of Lanka Electricity Company who were helped me to collect data from smart meters.

Finally, I would be thankful to my friends and family members including my wife and parents who always encourage and help me to complete this research.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

LV	Low Voltage
LECO	Lanka Electricity Company (Pvt.) Ltd.
CEB	Ceylon Electricity Board
GUI	Graphical User Interface
SSE	Sum of Square Error
P	Active Power
I	Current
V	Voltage
R	Resistance
S	Apparent Power
kVA	Kilo Volt Ampere
kW	kilo Watt
W	Watt

List of Appendices

CHAPTER 1

1. INTRODUCTION

In Sri Lanka, electricity customers can be divided in to two categories. One is bulk customers, which has dedicated transformer connection and other one is retail customers, which are fed by a common distribution transformer. Total electricity customer base of whole country in year 2017 is about 5.5 million and out of those, only about 20,000 customers have bulk connections. [Annual reports of LECO and CEB]

1.1. Background

Normally bulk customers are facilitated with an advanced electricity meter which has Remote Meter Reading facility and load profile recording facility. It is worthwhile to do such investment because, bulk customers are consuming considerable energy share and the number wise they are small percentage (0.36%) compared to total customer base in country. Therefore, extracting of load profile for bulk customers is not a problem. It is readily available in workstation servers.

Currently retail customers haven't load profile recording electricity meters. Retail customers are facilitated with basic energy meters which can read only the monthly energy consumption. It is not worthwhile to install an advanced electricity meter for each retail customer because, the number of customers is very large and the energy consumption of a single customer is very low compared to bulk customer.

But, estimation of load profiles of each retail customer connected to a distribution substation is very important for an electricity distribution utility. Distribution network planning cannot be done without knowing the loads of the customers. In current situation, low voltage distribution planning is being done by considering that all customer loads in specific feeder has same value and it is called After Diversity Maximum Demand.

$$After\ Diversity\ Maximum\ Demand = \frac{Peak\ load\ of\ transfomer\ in\ kVA}{No\ of\ customers\ connected\ to\ transformer}$$

Low voltage feeder power loss and voltage drop are calculated under the above assumption. But in actual situation, apparent power (kVA) demand of a customer has large variation with time of the day and different customers have different load profile patterns. Power loss of a feeder conductors ($I^2 * R$) is proportional to the square of the current in feeder. The variation of kVA loading of customers (Figure 1.1) affects the estimated power loss ($P = I^2 * R$) of feeder conductors in square form (Figure 1.2).

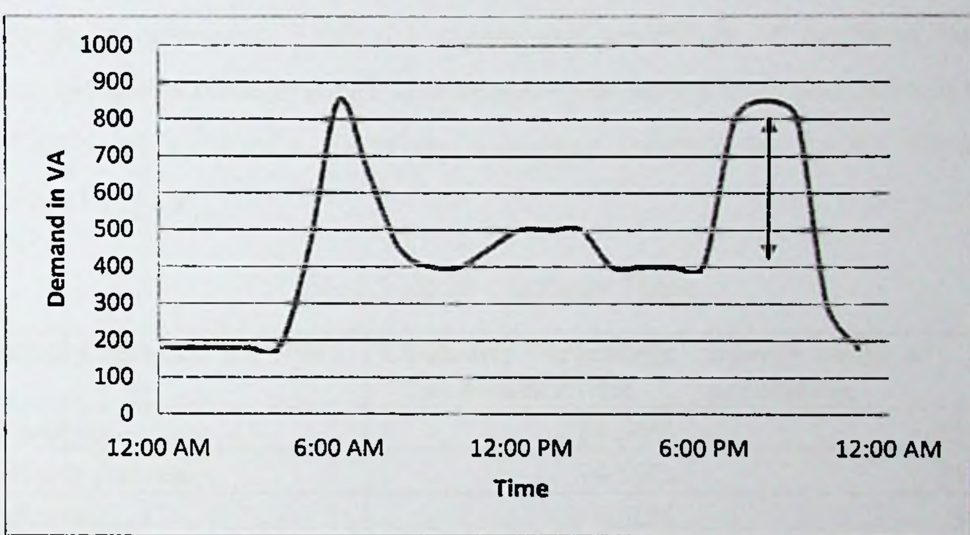


Figure 1.1 kVA Customer Demand Variation

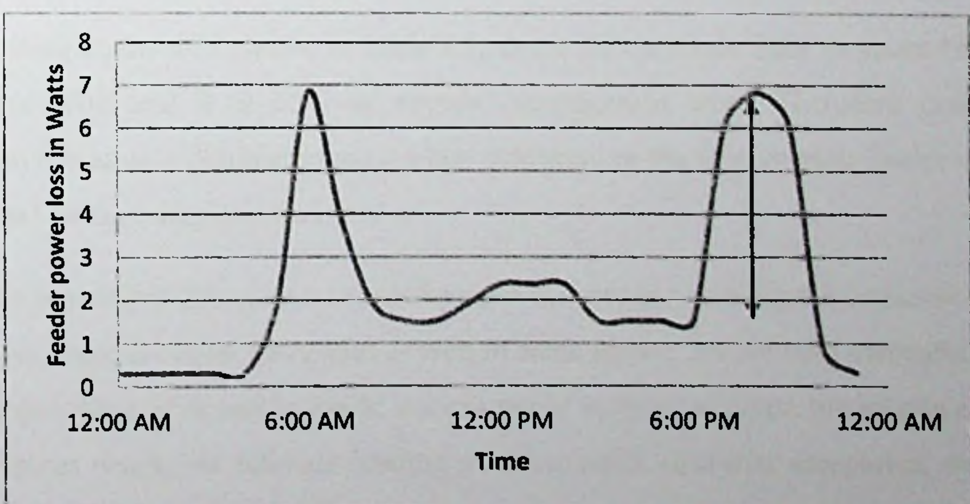


Figure 1.2 Feeder Conductor Power Loss

Therefore the variation of power loss is further increased according to this non linear relationship with customer demand. Assuming that, all customers has equal loading, will gives us inaccurate results for network power loss and for maximum voltage drop calculation and far deviated from actual situation. Therefore, time dependant customer loading and fare representation of scale of power demand of customer are very important in distribution network planning.

In Sri Lanka, retail customers are divided in to six categories those are fed by distribution transformers. Table 1.1 shows the percentage of customer base in different categories considering all distribution transformers as at year 2016 in whole area distributed by Lanka Electricity Company [records by Lanka Electricity Company (Pvt.) Ltd.]

Table 1.1 Retail Customer Types

Retail Customer Category	Customer Percentage as number wise	Energy usage as percentage
Domestic	84.86%	65%
General purpose	14.09%	31%
Industrial	0.50%	3%
Hotel	0.00%	0%
Government	0.07%	0%
Religious	0.48%	1%

According to the data shown in table 1.1, domestic customer base is about 85% in number wise and it is 65% as energy consumption wise. Therefore domestic customer load is a dominant factor when determining the low voltage feeder power loss and voltage drop calculations.

Feeder power loss calculation cannot be done by only considering the domestic loads because, there are other categories as well in same feeder. But without estimating this dominant share of domestic loads, we can never achieve accurate feeder power loss calculation result. To estimate loading of other retail customer categories, another study is needed. Otherwise, installation of load profile recording meters for other customers is worthwhile because the number wise percentage is 15% and energy consumption wise percentage is 35% of total load.

The topic of this research is "Defining of Normalized Load Profile Curves for Domestic Customer Groups to Estimate Feeder Power loss". The research part of this topic is "Defining of normalized load profile curves for domestic customer groups". "Estimation of feeder power loss" is the application of research outcomes for a useful purpose in distribution utility. In this research, the domestic customers will be divided into number of groups depending on their load pattern and method of defining of load profile for a new customer is described in problem statement and methodology.

Estimation of load profiles for domestic customers has multiple purposes such as tariff design of domestic customers, differentiated customer service, identifying network stress points, identifying high demand periods of separate feeders, voltage drop calculation, and much more. Estimation of time dependant feeder power loss is only one application of defining domestic customer load profiles.

1.2. Problem statement

When we considering about energy loss from power generation point to customer end, low voltage distribution energy loss is the dominant. So, calculation of accurate figure for low voltage network energy loss is very important.

A simple Low voltage feeder of a distribution transformer can be shown as in figure 1.3. In Sri Lanka situation, the retail customers are facilitated with basic accumulated energy reading electricity meters at customer end.

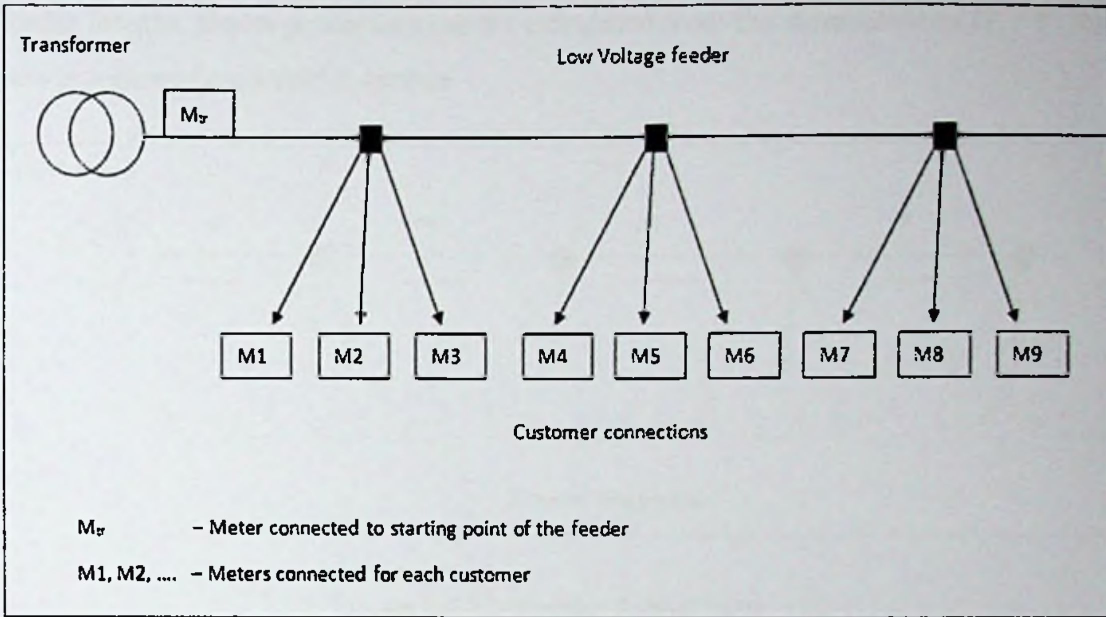


Figure 1.3 LV Feeder Metering Points

Theoretically, low voltage feeder energy loss in specified time period should be equal to energy reading of M_{tr} – (M1+M2+M3+.....). But practically those meters have +/- 1% error. Therefore power loss calculation by reducing those meters readings has unpredictable error.

As an example:

If actual energy transferred through meter M_{tr} is 104 units, actual loss in the feeder is 4 units and actual total energy delivered to consumers is 100 units. Actual energy loss should be nearly 4% In worst case, the energy reading of M_{tr} may be 103 units and summation of energy readings at consumer points may be 101 units due to +/-1% error of meters. Then calculated energy loss percentage will be nearly 2%. Therefore energy loss calculation by reducing the energy meter readings may give us about 50% error. So it is clear that this is not a correct approach to estimate feeder energy loss. Time dependant power loss is not feasible at all with this basic energy meters.

More accurate way to calculate time dependant power loss is using estimated load profiles of the consumers and following 'forward backward sweep method' to calculate currents of each feeder sections. Then, by using conductor resistances and feeder lengths, feeder power loss can be calculated from the summation of ($P = I^2 * R$) power losses of each feeder section.

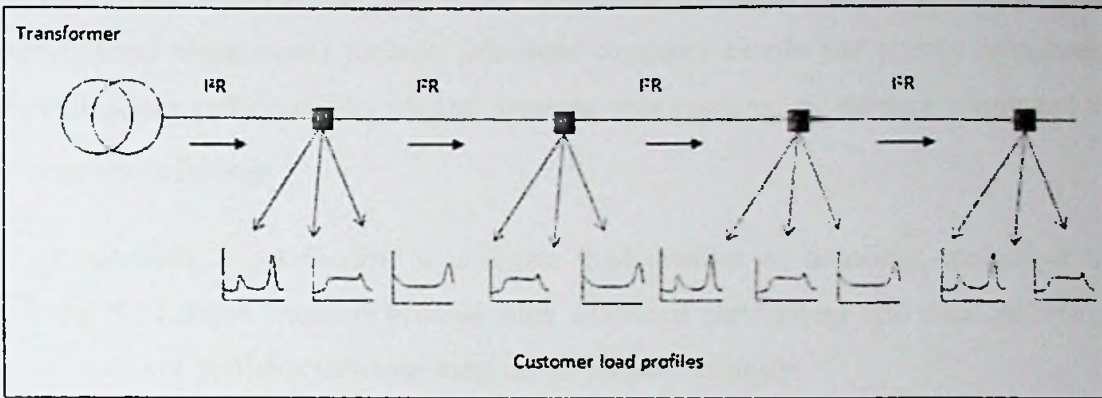


Figure 1.4 LV Feeder Power Loss

Algorithm for calculation of feeder power loss is proposed at 6th chapter of this thesis and that can be implemented transformer-wise in power distribution utility if they have the network information like conductor sizes, conductor lengths etc...

The major problem for this power loss calculation approach is unavailability of customer load profiles. There are three options to obtain load profiles of domestic customers.

1. Installation of meters with load profile recording facility

The domestic customer base is very large (about 5 million] in number wise and installation of load profile recording meters for all the customers is not practical and very expensive project. In rural areas, household energy consumption is very low. Hence installation of advanced energy meter for those customers is not worthwhile. Therefore the only way to obtain load profiles of domestic customers is to go for an estimation method.

2. Big data analytics

A definition for big data analytics was found from website 'Techtarget Network' it says, "Big data analytics is the process of examining large and varied data sets -- i.e., big data -- to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful information that can help organizations to make more-informed business decisions. That encompasses a mix of semi-structured and unstructured data -- for example, internet click stream data, web server logs, social media content, text from customer emails and survey responses, mobile-phone call-detail records and machine data captured by sensors connected to the internet of things."

This approach is not feasible to estimate load profiles of domestic customers in current Sri Lankan situation because such advanced networking and data gathering means are not available covering most of the areas of country.

3. Collecting information from customers by a questionnaire

The feasible approach to collect information in current situation is to prepare simple questionnaire and get it filled by customer. Meter reading officers are coming to customer premises in every month to get meter readings. Power distribution utility can distribute this questionnaire though meter reading officers and collect those filled forms in consecutive month.

Methodology part of this thesis presents a software application that can be installed in a tablet or any other mobile device to collect customer information and upload to server. That can also be done by meter reading officers or separate survey team as special project.

In this research the customer information related to family member composition, electric equipments used and the time/nature of using those equipments will be collected. Average monthly energy consumption of each customer will be obtained from past data of electricity bill preparing program.

1.3. Objectives of study

As described in problem statement, the problem is unavailability of load profiles of domestic customers to estimate time dependant feeder power loss. This study has two objectives,

1. Defining of load profile curve for any customer depending on his information such as family member composition, electric equipments used, the time/nature of using those equipments and system data of past average monthly energy consumption of that customer.
2. Development of methodology to estimate time dependant low voltage feeder power loss (power losses in conductors) for distribution transformers.

1.4. Motivation

This research enables to obtain domestic customer load profiles without installing load profile recording meters for every customer. Low voltage feeder power loss calculation is only one use of domestic customer load profiles estimation. There are other uses that are very important for an electricity distribution utility such as,

- Low voltage substation and network planning
- To identify network stress points
- To identify high demand periods of separate feeders
- Tariff design
- Study the demand elasticity with tariff revisions
- To design market strategies for utilities
- Differentiated services for customers

If updated information about the customers and relevant load profiles are available in the databases, above activities can be implemented by the utility.

1.5. Methodology

Major activities of this research can be ordered as follows.

1. Load profile data collection of randomly selected customers in a selected area

For the first time in Sri Lanka, smart meters were installed for domestic customers with load profile recording facility in year 2016. This pilot project was done by Lanka Electricity Company (Pvt.) Ltd. in area of Sri Jayawardhanapura Kotte for 3200 numbers of domestic customers. For this research, 81 numbers of customers were selected and load profiles were collected for 10 week days for each customer. Those 10 load profiles were averaged to obtain one representative load profile for a customer to be more accurate.

2. Clustering of load profiles to group similar load patterns

The purpose of clustering is to group similar load patterns. Since, distance based clustering algorithms (k-means algorithm in MATLAB) are used, all 81 numbers of load profiles were normalized to equal area under each curve. Then only the pattern matters for clustering and effect of magnitude of the load profile is removed. Thereby 8 groups of customers with differentiated load profile shapes were identified.

3. Do surveying and collecting information of those customers

Then all the customers in 8 groups were interviewed and information such as family member composition, electric equipment information and salary scale of family were collected.

4. Analysis of customer information and load profile to identify relationships

Collected information was analyzed to find out similarities in same group with relevant load profile pattern. Those 8 different representative load profiles were named as follows for identification purpose.

- Time restricted job holders
- Students
- Unemployed/retired/self employed people (at home)
- Flexible time outside job holders
- Evening load consumers
- Noon load consumers
- Night time load consumers
- Even load consumers

Now, there are 8 numbers of differentiated shape load profile curves and identification name for that cluster.

5. Development of algorithm to find out unknown load profiles from customer information

Final objective is to select the best match load profile curve out of 8 above mentioned curves according to customer information, for a customer which is not having load profile recording meter.

MATLAB graphical user interface was developed to input customer family information and electric equipment usage information. Then that program will output the best match representative load profile curve out of 8 curves for that specific customer.

6. Application of load profiles to estimate time dependant feeder power loss

Example is presented at the end of this thesis for calculation of feeder power loss. Up to now, selection of best match representative curve according to customer information is discussed. But, that selected curve is a normalized curve so that area under the curve is 1kWh. We can obtain average monthly kWh consumption of the customer by bill preparing program in distribution utility. Thereby, daily kWh consumption can be deduced. By multiplying representative per kWh curve by actual daily kWh consumption, we can obtain actual kW load profile of that customer. That kW load profile will be divided by assumed power factor to obtain required kVA load profile.

Network data such as conductor lengths and sizes are available in Geographical Information System of utility or otherwise those information need to be updated, voltage profile of feeder starting point is captured by the smart meter installed at distribution transformer. Feeder section currents are calculated by 'Forward Backward Sweep Method' and thereby the time dependant feeder power loss is calculated in 24 hour time frame of the day.

CHAPTER 2

2. LITERATURE REVIEW

Most of the international research papers are describing load profile clustering methods, but not described the application of the clusters for a useful purpose in distribution systems. The papers just mentioned the various uses of grouping of the domestic customers, [2] [3] but they have not presented a method of implementing any of those, because, that is job of electricity distribution utility.

There were some researches done by using smart meters for domestic customers. [1] They have used kWh consumption, maximum demand and minimum demand as attributes for clustering. This research is totally different approach from those.

In this research, I have presented research part as well as application part, because I have done this research conjunction with one of electricity distribution utility in Sri Lanka called "Lanka Electricity Company (Pvt.) Ltd". They have practiced a domestic customer classification method about 10 years back and definition of representative load profile curves for those groups.

They have divided domestic customers in to groups according to their monthly energy consumption as below.

- Group 1 – below 30 units per month
- Group 2 – 30-60 units per month
- Group 3 – 60-90 units per month
- Group 4 – 90-150 units per month etc...

Then set of load profile recording meters were connected to each customer in one group and obtained load profiles. Those load profiles were used to calculate average load profile curve for that specific group. Drawback of this approach is that the patterns of load profiles of different customers in same group may be different. So they are averaging load patterns with large variation in shape.

Figure 2.1 is prepared to shows the lapse of that old method.

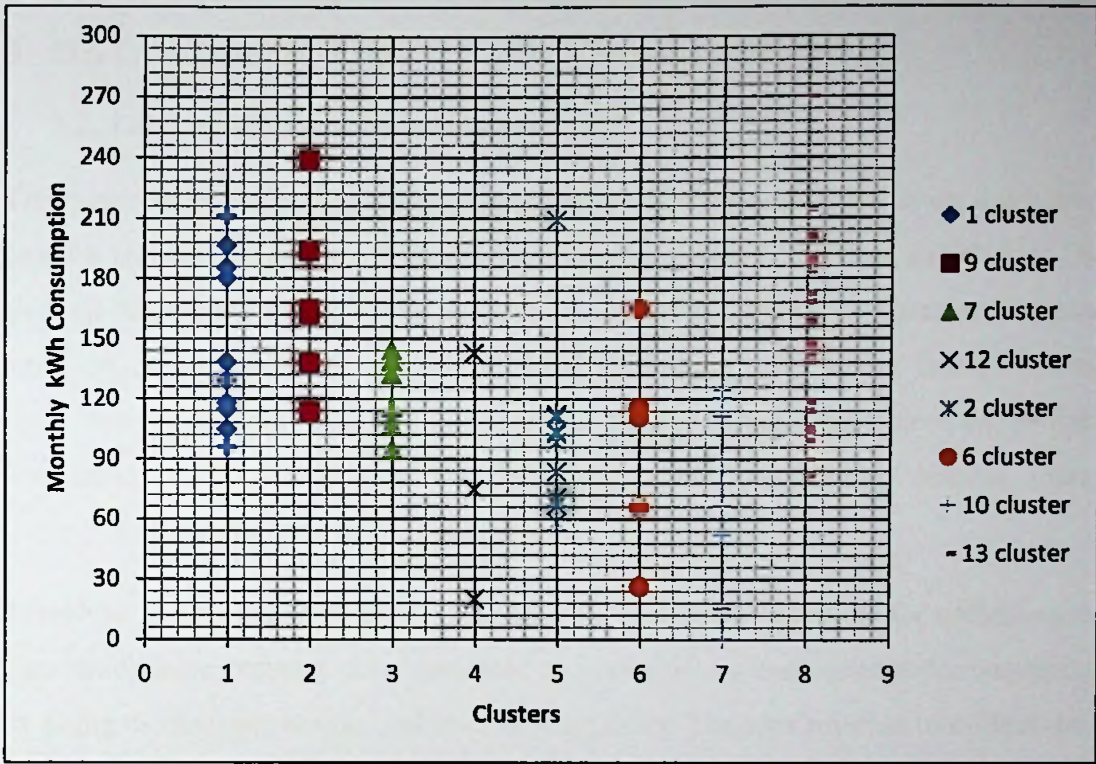


Figure 2.1 Monthly Energy Consumption Distribution

This figure was deduced after classification of customers according to the approach presented in this research. In my approach I have divided customers in to 8 groups according to the similarities in load profile pattern. If we consider cluster 1 in figure 2.1, the monthly energy consumption of customers in that group has a large range between 90-210 units. If we consider cluster 6, the monthly energy consumption variation of customers is between 30-180 units.

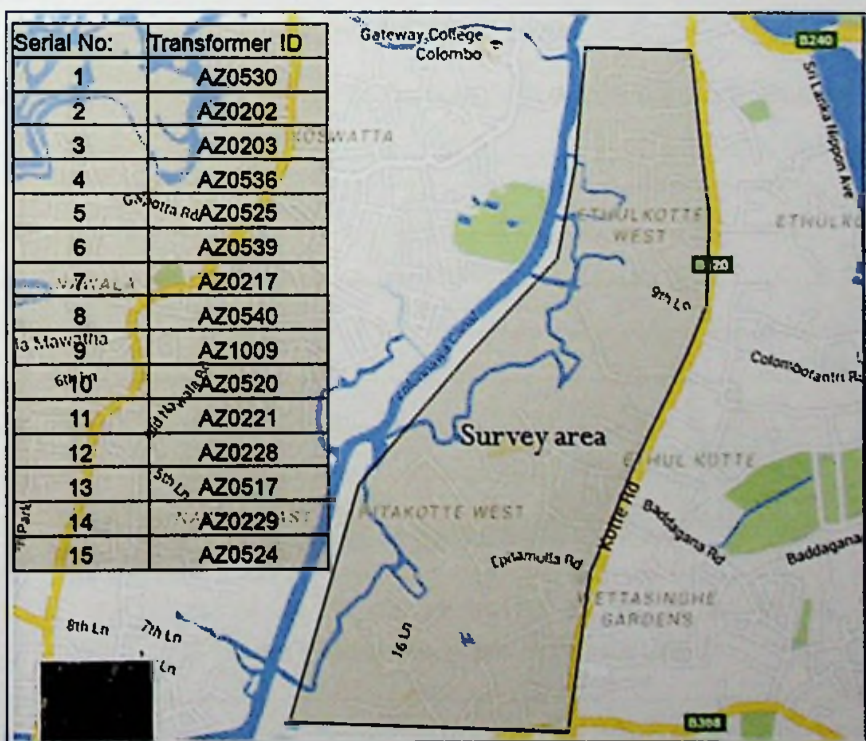
If we analyze in other way, in monthly energy consumption range of 60-90 units, there are customers in 5 numbers of clusters. In monthly energy consumption range 90-120, there are customers in all the clusters. This reveals that, grouping of customers considering monthly energy consumption range is not a right approach to define representative load profile curves for customer groups. But 10 years back, that approach is used because of unavailability of installed smart meters for domestic customers.

3. DATA COLLECTION

3.1. Load profile from smart meters

The customer classification method discussed in this research is based on daily load profiles patterns. There are 3200 numbers of smart meters installed in 2016 of the area Sri Jayawardenapura Kotte by Lanka Electricity Company. There were some revisions of the firmware versions of smart meters and only about latest 1200 of them were enabled for load profile recording facility. Since smart metering project was done in Sri Lanka for the first time, there were some remote meter communication problems as well at that time.

Therefore, in this research, I will consider 81 numbers of customers for collecting of data which were properly communicated and considering convenience for surveying by going to customer houses and interviewing them. The area selected to collect data is shown in figure 3.1.



Serial No:	Transformer ID
1	AZ0530
2	AZ0202
3	AZ0203
4	AZ0536
5	AZ0525
6	AZ0539
7	AZ0217
8	AZ0540
9	AZ1009
10	AZ0520
11	AZ0221
12	AZ0228
13	AZ0517
14	AZ0229
15	AZ0524

Figure 3.1 Survey Area

81 numbers of customers were distributed among 15 numbers of distribution transformers as mentioned in figure 3.1. Customers in this area have a large range of electricity consumption variation from about 50 units to 500 units per month. This is not a very rural area as well as not highly urbanized area. Living pattern of customers has a large variation and occupations are also different. Figure 3.2 shows a picture of smart meters that are used for data collection.

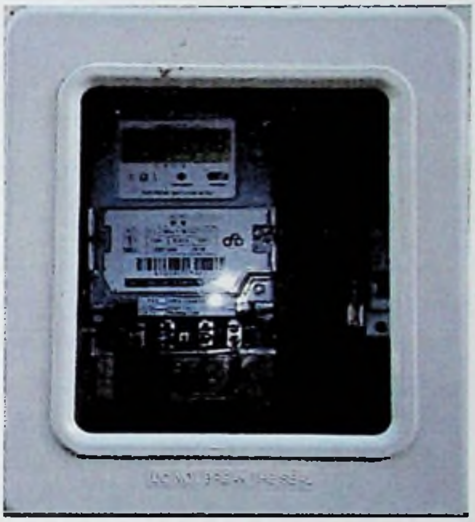


Figure 3.2 Smart Meter

Figure 3.3 shows extracted load profile curves for 10 days for a one customer.

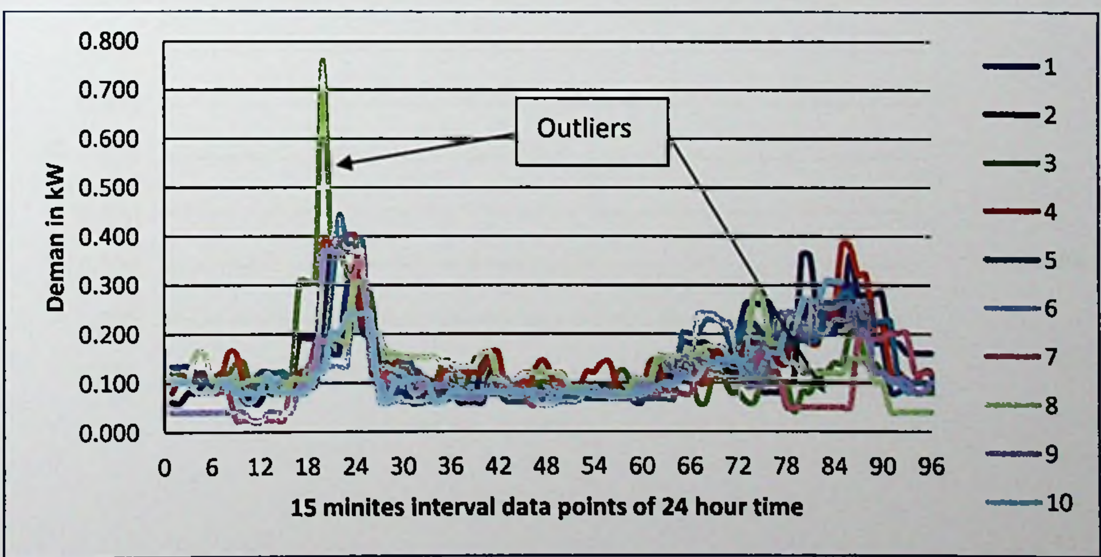


Figure 3.3 Customer Load profiles of Several Days

Some abnormal behavior of load patterns may happen unexpectedly due to special occasions. Those are called outliers and those are removed when getting an average daily load profile for a customer as in figure 3.4.

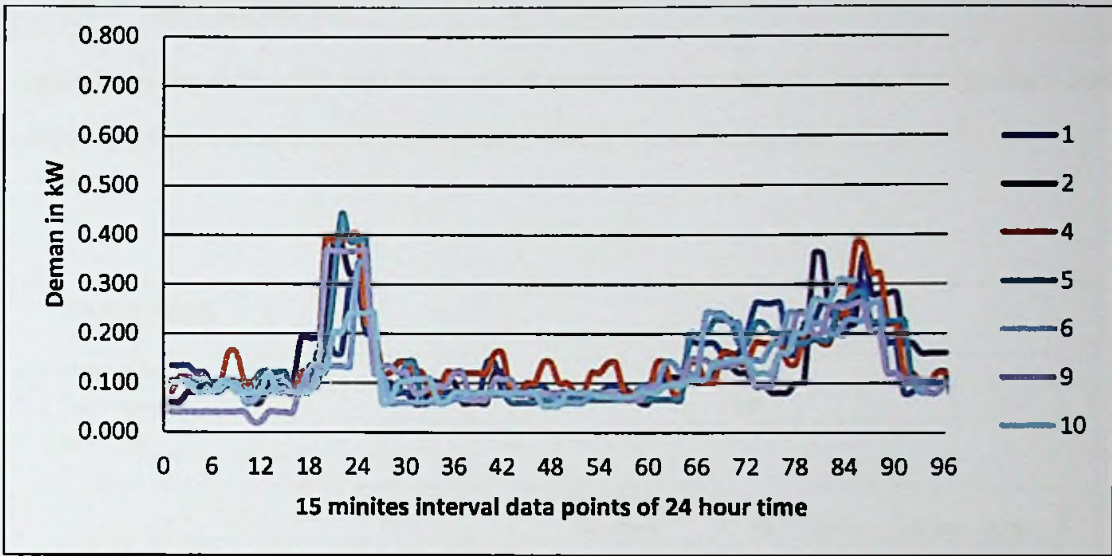


Figure 3.4 After Removing Outliers

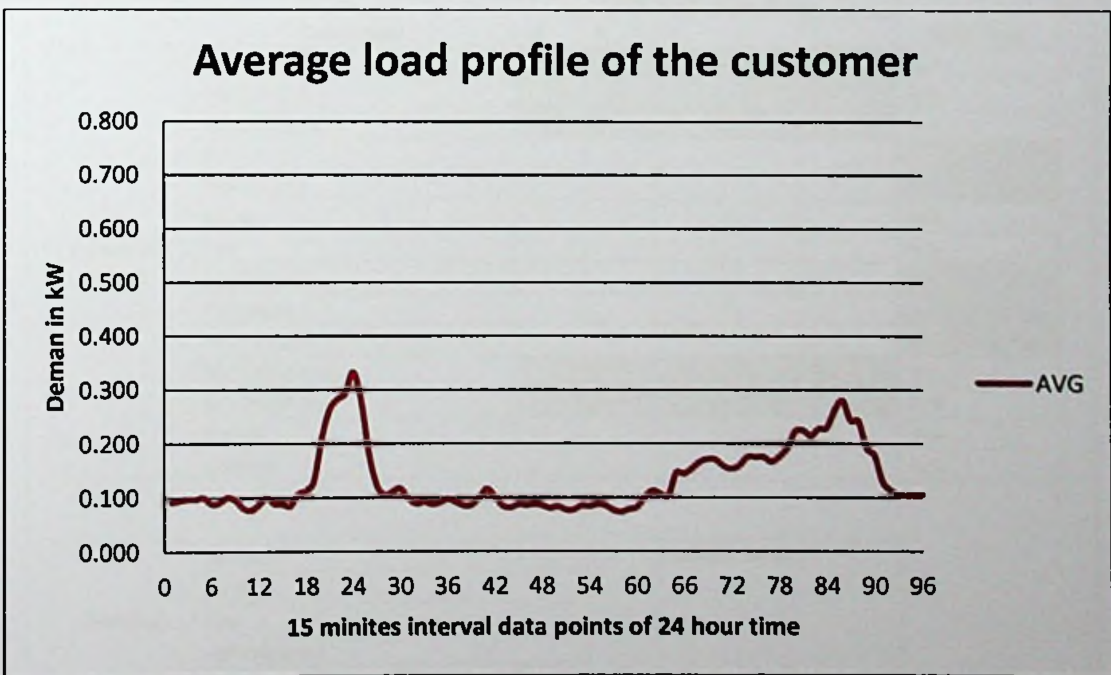


Figure 3.5 Daily Average Load Profile of a Customer

Load profiles of 81 numbers customers were analyzed and average load profile curves were obtained in this way for all customers.

3.2. Customer Survey

Apart from the data extracted from smart meters, 81 customers were interviewed and gathered information according to survey sheet shown in figure 3.6.

Survey Sheet

Account No:	
Transformer:	
Serial No:	

Units:	
Bill :	

		How many members
Member composition	Students	
	Restricted time job holder	
	Flexible time out side job holder	
	Unemployed people	
	Retired people	

M	Morn.	05:00 - 10:00
N	Noon	10:00 - 13.00
A	Af.Noon	13:00 - 18:00
E	Eve.	18:00 - 22:00
T	Night	22:00 - 05:00

		Used time/ How many					
	Equipment	M	N	A	E	T	Size/ Type
Equipment usage	Bulbs						N/A
	Fan (Ceiling)						N/A
	Fan (Pedestal)						N/A
	TV						Inch
	Rice cooker						kg
	Kettle						N/A
	Iron						N/A
	washing mc.						kg
	Computer						Desk Top/Lap
	Refrigerator						feet
	AC (12000 Btu)						N/A
	AC (24000 Btu)						N/A
	Geyser						N/A
	blender						N/A

Dwelling	Rooms	
	Stories	
	Size	
	Lighting level	
	Closed / Open	

Special needs:

Figure 3.6 Survey Sheet

The top information of figure 3.6, Account number is the number given by the utility for a specific customer as unique identification. Serial number is the number mentioned in smart meter, because that is required when mapping this information with smart meter load profile data. Monthly energy consumption units and monthly bill was noted as reference, but past data can be collected from billing system database.

Members of the family are divided into 5 categories considering their energy usage behavior and time. When students are in a family, they will get up early morning and come back in evening. Students may have a habit of studying in morning or night. Restricted time job holders may be working for government or private sector, but they are leaving home in morning at same time and coming back in evening at their usual time. Flexible time outside job holders may leave home at late morning or at another time, however they are doing their jobs in outside. When unemployed people, self employed people and retired people are at home, day time energy using activities may happen.

These behaviors of different types of members affect daily energy consumption pattern and it is reflected in load profile curves captured by smart meters. Conducting of survey to collect these information and load profiles collected from smart meters are needed to analyze the relationship between member composition of family and load profile.

The member composition alone will not define the energy consumption pattern of the day, because families with same type of members may use different types of electric equipments. Therefore the list of electric equipment used in that house is needed. And the size of equipment is needed for some items such as air conditioners, refrigerators, rice cooker etc...

If the some families even have same member composition, same electric equipments, the energy consumption pattern may differ according to their habit of using those electric equipments. Some people used to iron cloths in morning; some are at evening or night. Some people use washing machine in morning, some are in evening. There are special cases, the families which has babies, they use washing machine at noon in everyday. Some people turn on T.V. in early morning to listen religious blessings. Therefore the habit of using electric equipment is very important to determine energy consumption pattern.

Special electricity needs are also important if some members are using electricity for self employments, such as sewing machines, cooking meals as business etc... Lighting level of dwelling and air circulation ability of the dwelling may affect day time lighting load and usage of fans.

The survey sheet shown in figure 3.6 was prepared considering these things which have a possibility to affect the energy consumption pattern. Some of these may not considerable effect and some have a great effect for load variation.

The survey sheet was prepared with simple questions where anyone can answer. The information such as monthly salary, and other deep personal information that are difficult to get a reliable answer, are not included in survey sheet because, the people are not providing trustful information for those. Then whole study will depend on non-reliable data.

4. DEFINING OF DOMESTIC CUSTOMER GROUPS

Averaged load profile curves of 81 numbers of domestic customers have different shapes. Daily energy consumption is represented by the area under the 24 hour time kW load profile curve and there is a large variation of the magnitude of area. This study is based on categorization of domestic customers according to the shape of their load profile. Finally, the customer load profiles in same group should be same in shape, but the magnitude may be different.

In figure 4.1, both load profiles have same shape, but magnitude is different.

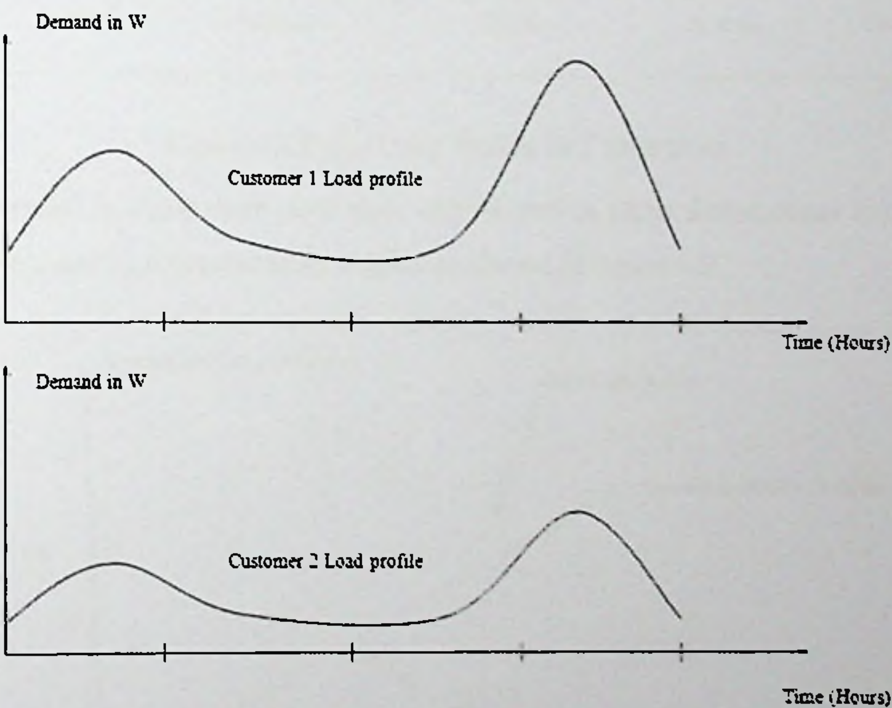


Figure 4.1 Sample Load Profiles

In this study, these customer 1 and customer 2 should be in same group, because the load profile pattern is same. Therefore, a method is needed for numerical representation to identify difference or similarity of the load profile patterns to cluster customers with same load profile curve shapes.

4.1. Numerical representation to identify difference of load profiles

If we consider a load profile with only three time slots of the day, it can be shown as in figure 4.2.

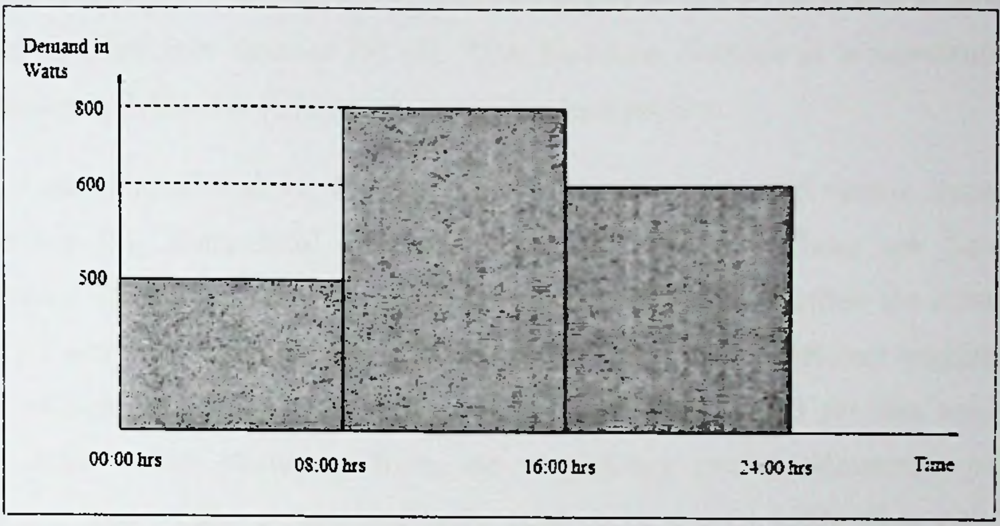


Figure 4.2 Daily Load Profile as 3 time slots

If the demand in these three time slots represented in three dimensional space, this load profile can be represented by a point as shown in figure 4.3.

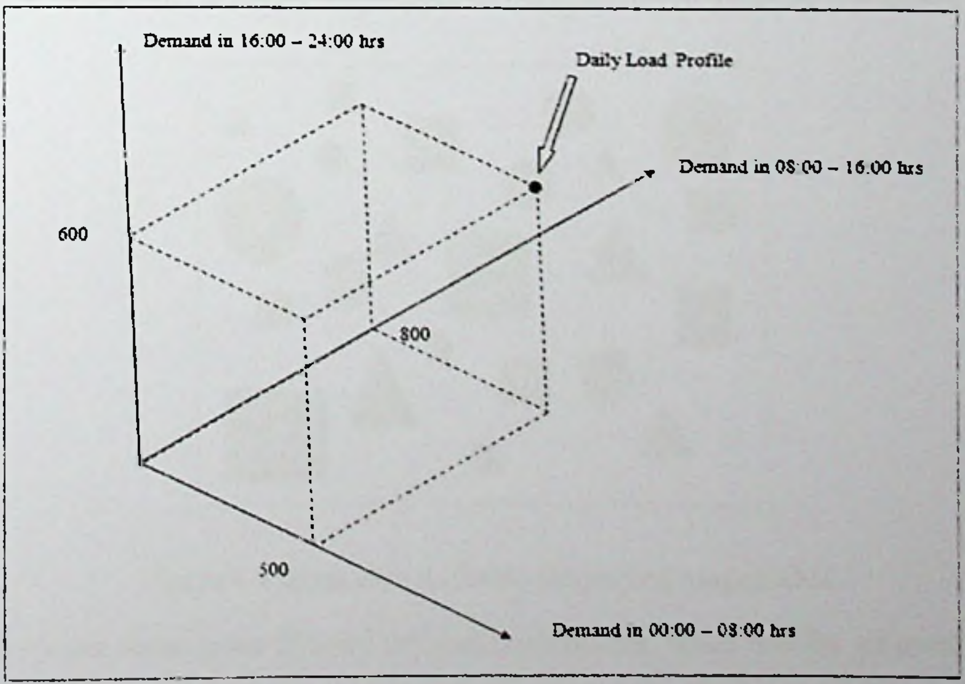


Figure 4.3 Load Profile in 3D space

Likewise if, the load profile has 96 data points in 15 minutes intervals of 24 hour domain, the customer load profile can be represented as a point in 96-dimensional hypothetical space. [2] When that point moves to another place, the shape of load profile is changed. The distance between two points in this 96 dimensional space is called as Euclidean distance [8] [3]. That Euclidean distance is a representative parameter to define the difference between two load profiles.

There are standard methods for load profile clustering such as K-means clustering algorithm [8], Hierarchical clustering algorithm [8] etc... Those are Squared Euclidean Distance based clustering methods. If those load profiles are clustered directly without normalizing, the curves with same shape, but different magnitudes will belongs to different groups. Therefore, the customer load profiles must be normalized before clustering. Here, the normalizing means, adjustment of all customer load profiles so that area under each curve to be 1 kWh. Then only the difference of shape is represented by Euclidean distance and the effect of magnitude is removed.

To understand the concept, let's consider a simple example as figure 4.4 represents different shapes and different magnitudes of load profiles of different customers.

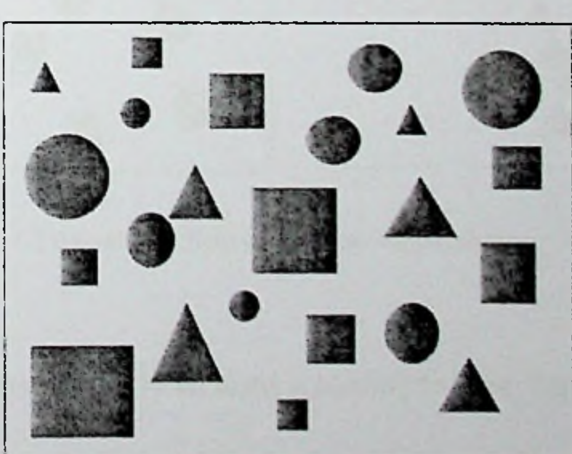


Figure 4.4 Icons with different shapes and magnitudes

If we cluster these icons directly without normalizing, those will be grouped as in figure 4.5.

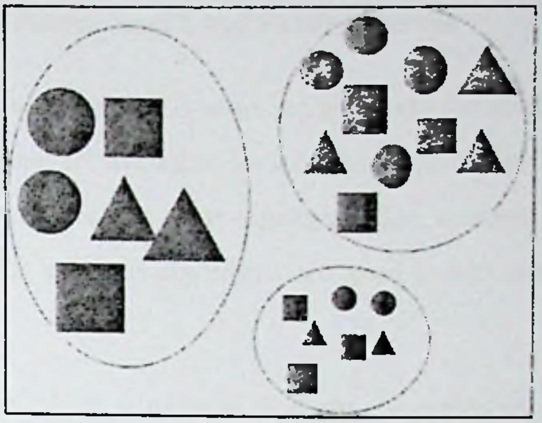


Figure 4.5 Clustered without Normalizing

In figure 4.5, the purpose of differentiating icons according to shapes is not succeeded. Therefore we have to normalized those icons before clustering. After clustering the normalized icons, those will be clustered as in figure 4.6.

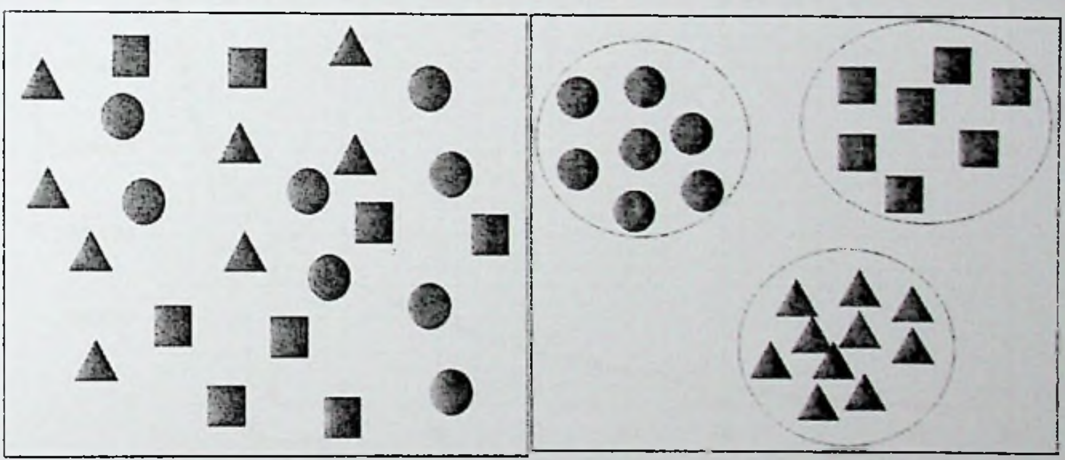


Figure 4.6 Clustered After Normalizing

Figure 4.6 shows correctly clustered icons according to their shapes.

Therefore 81 numbers of customer load profiles were normalized so that, area under each curve to be 1 kWh. Averaged raw load profile data of 81 customers are attached in Appendix G and 81 number of normalized load profile data in Appendix H as soft copy.

4.2. Clustering of Normalized Load Profile Curves

K- Means clustering algorithm [8] is used in this study for load profile clustering. In built MATLAB functions are available for k-means clustering. There are 81 numbers of normalized load profile curves to be clustered. But, number of clusters in that set of data is unknown. For k-means algorithm, we need to define the number of cluster before clustering.

Knee point analysis was done in order to find out the optimum number of clusters in that data set. Figure 4.8 shows the decreasing of sum of square error (SSE) [5] verses number of clusters.

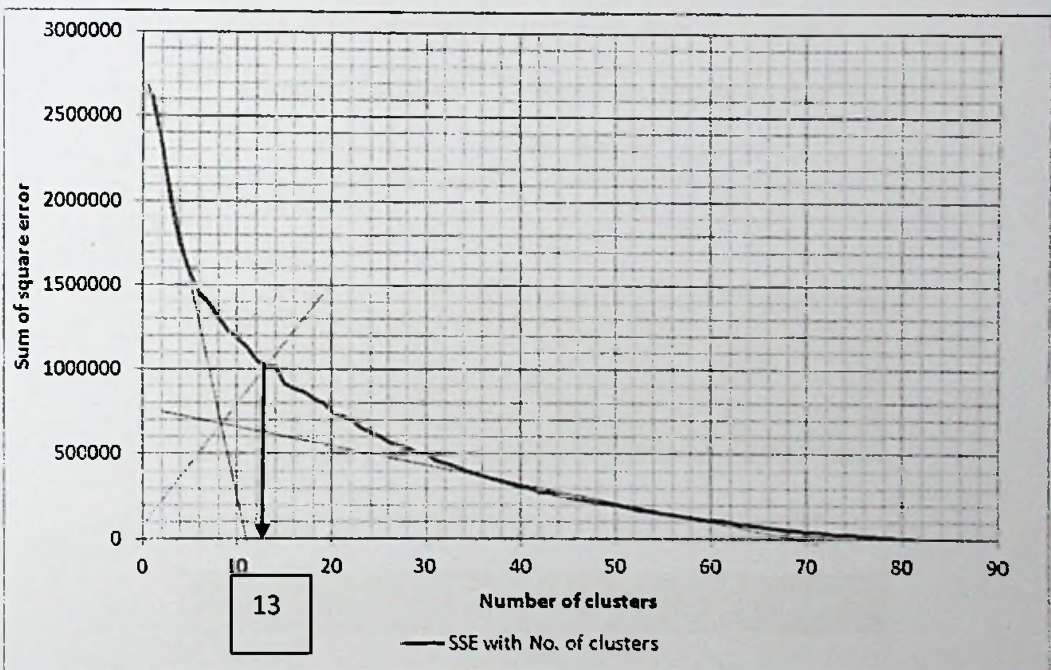


Figure 4.7 Optimum Number of Clusters

First, 81 numbers of load profiles were clustered in to 1 cluster and sum of square error was calculated. In second step, those 81 numbers of load profiles were clustered in to 2 clusteres and calculated the sum of square error. Likewise, the data set was clustered up to 81numbers, and calculated SSE and plotted 'SSE' verses 'number of clusters' as in figure 4.8. If the numbers of cluster is equal to numbers of load profiles, SSE will be zero.

To find out optimum number of clusters, intersection of two trend lines for first 6 data points and last 50 data points is considered. Angle bisector is extended from trend line intersection point up to the SSE curve to find out knee point of the curve. Knee point is found at 13 numbers of clusters in x-axis, hence the optimum numbers of cluster is found as 13.

Inbuilt function for k-means clustering was used to cluster 81 numbers of load profiles in to 13 numbers of clusters. The result was indicated that the number of load profiles belongs to each cluster as in table 4.1.

Table 4.1 Customer Distribution Among Clusters

Cluster Number	**Number of load profiles in cluster**
Cluster 1	10
Cluster 2	12
Cluster 3	1
Cluster 4	2
Cluster 5	1
Cluster 6	5
Cluster 7	6
Cluster 8	1
Cluster 9	6
Cluster 10	11
Cluster 11	1
Cluster 12	3
Cluster 13	22
Total	**81**

According to the clustering results, cluster 3, 4, 5, 8, 11 has only one or two load profiles and those clusters cannot be considered as rich clusters. In this analysis, only the clusters having at least 3 load profiles were considered as rich clusters because the total number of load profiles is limited to 81.

Therefore 8 numbers of domestic customer groups were identified according to the clustering of normalized load profiles curves as in table 4.2.

Table 4.2 Set of Rich Clusters

Cluster Number	**Number of load profiles in cluster**
Cluster 1	10
Cluster 2	12
Cluster 6	5
Cluster 7	6
Cluster 9	6
Cluster 10	11
Cluster 12	3
Cluster 13	22

4.3. Defining of representative normalized load profile for customer groups

Each group has number of customer load profiles. Those are slightly different but in same cluster. The most representative load profile for a group is given by the cluster centroid of that cluster. Therefore to define a representative normalized load profile for each group; all load profiles in that group should be averaged [8]. Deduced representative curve is also a normalized curve so that, the area under the curve is 1kWh.

Figure 4.9 shows the normalized load profiles in cluster 1.

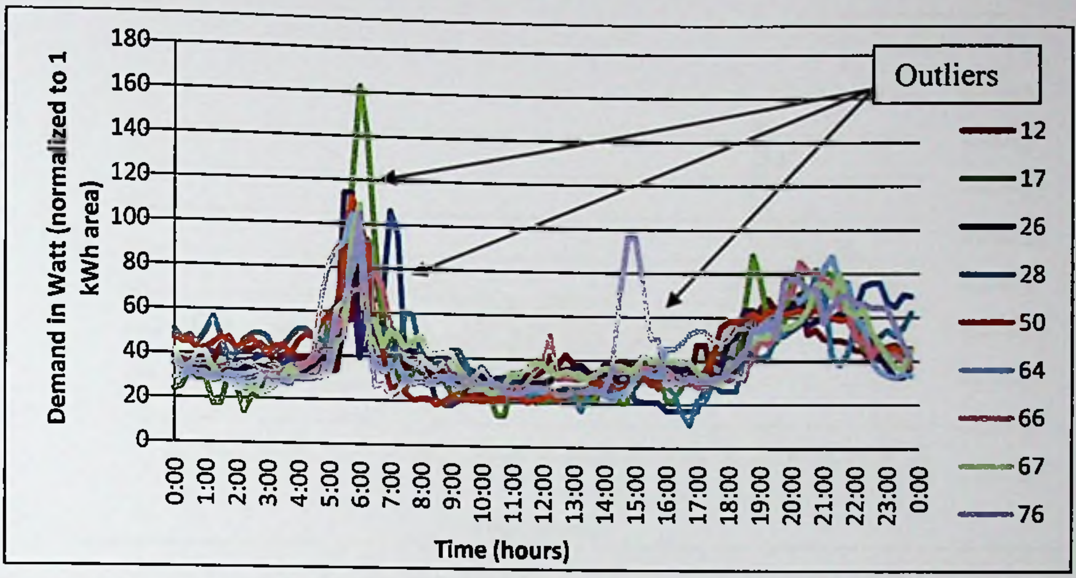


Figure 4.8 Cluster 1

As shown in figure ##, the curve number 3, 17, 76 are outliers because those are deviated from common profile of that group. Therefore those curves were not considered when obtaining representative normalized load profile for cluster 1. Representative normalized curve for cluster 1 is shown in figure 4.10. This customer group has a morning load peak than evening peak.

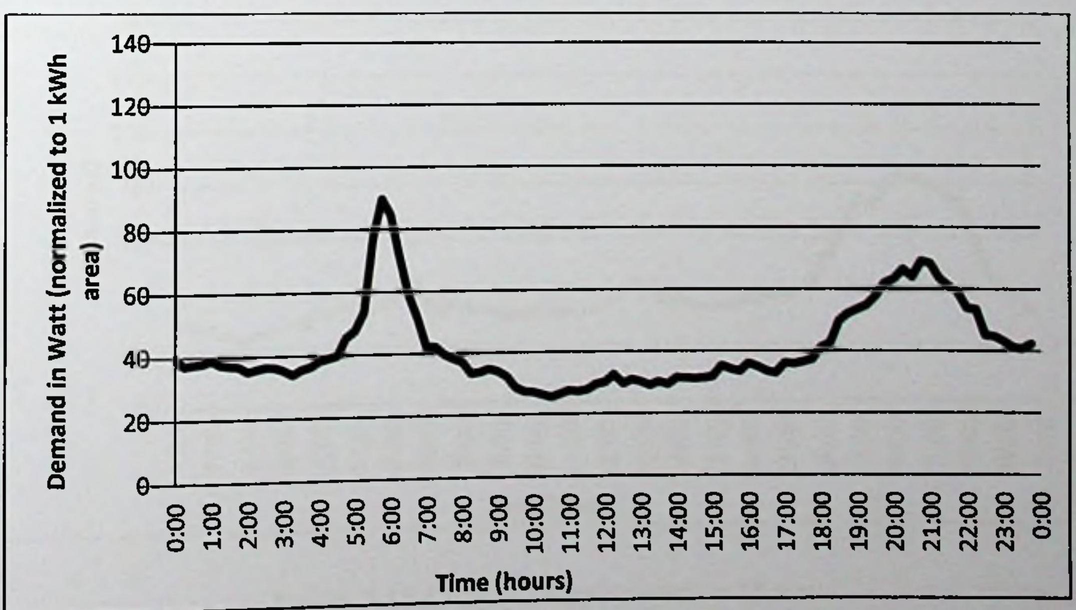


Figure 4.9 Cluster 1 Representative Curve

Figure 4.11 shows the normalized load profiles in cluster 2.

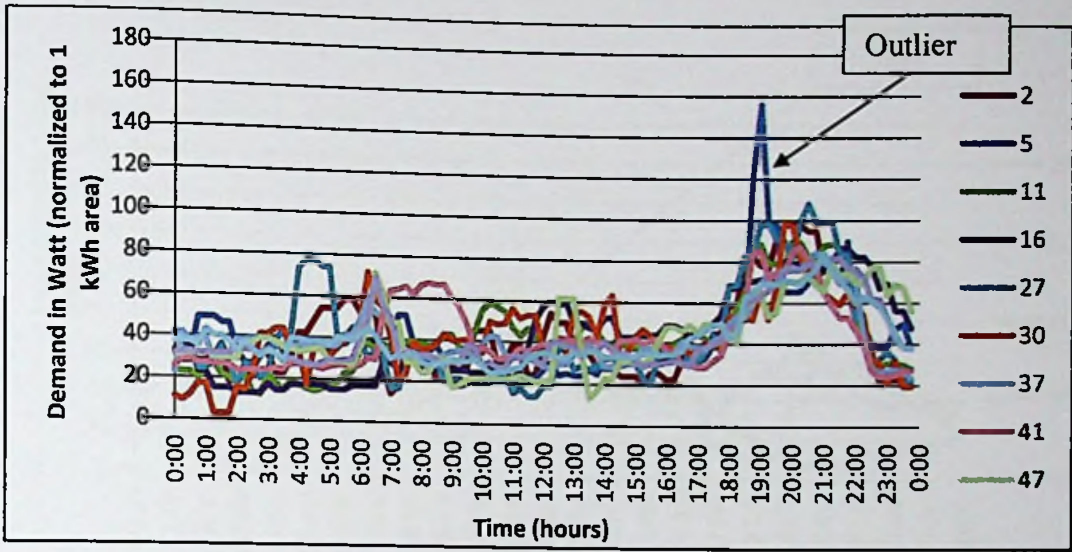


Figure 4.10 Cluster 2

As shown in figure 4.11, the curve number 5 is an outlier because that is deviated from common profile of that group. Therefore that curve was not considered when obtaining representative normalized load profile for cluster 2. Representative normalized curve for cluster 2 is shown in figure 4.12. This customer group shows even load at all the time except evening.

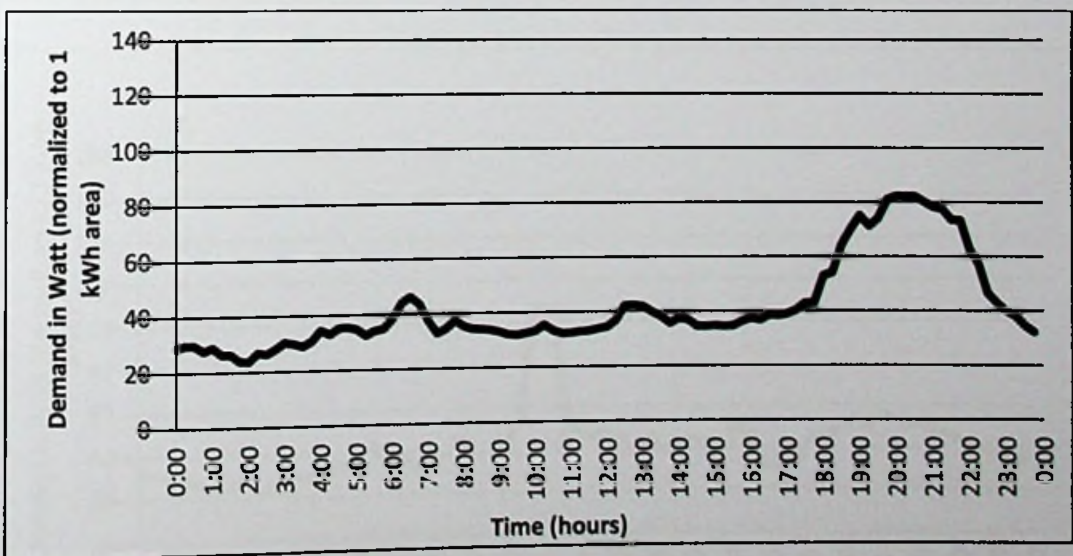


Figure 4.11 Cluster 2 Representative Curve

Figure 4.12 shows the normalized load profiles in cluster 2.

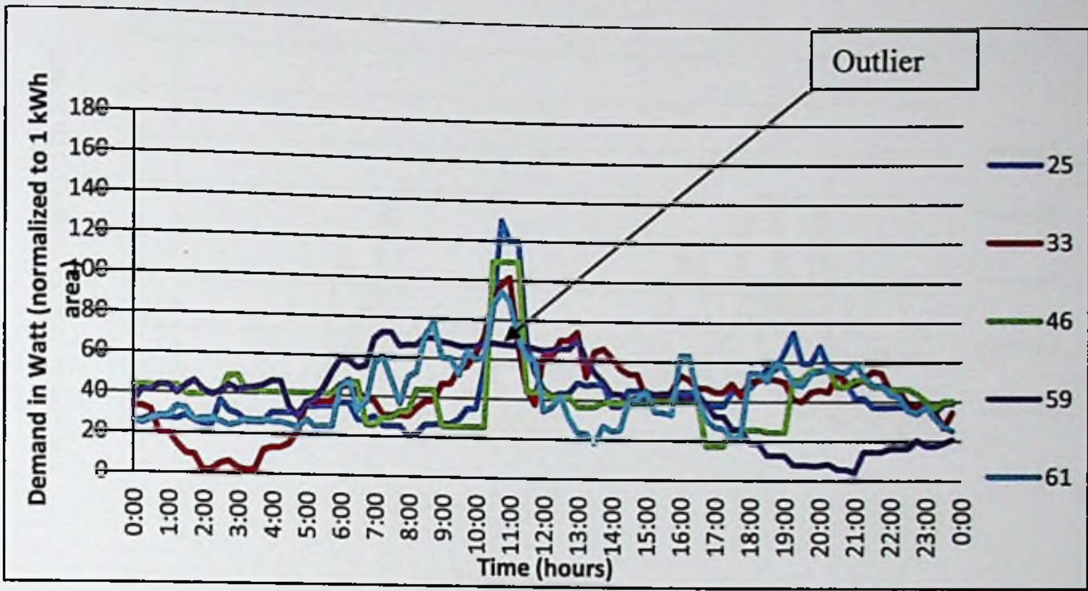


Figure 4.12 Cluster 6

As shown in figure 4.13, the curve number 59 is an outlier because that is deviated from common profile of that group. Therefore that curve was not considered when obtaining representative normalized load profile for cluster 6. Representative normalized curve for cluster 6 is shown in figure 4.14. The load peaking of this customer group is at noon.

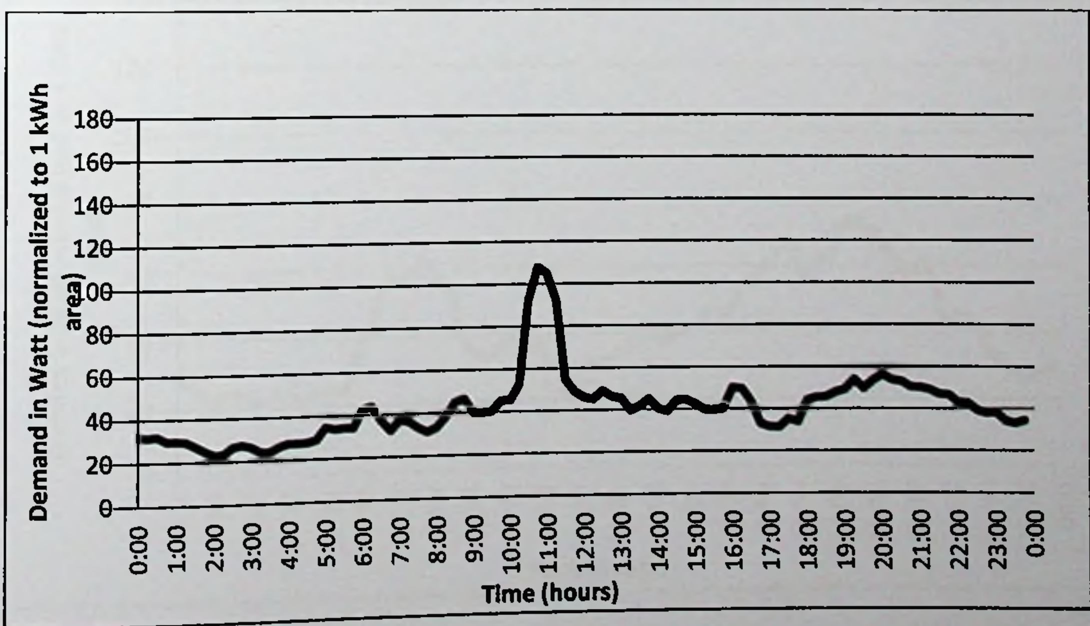


Figure 4.13 Cluster 6 Representative Curves

Figure 4.15 shows the normalized load profiles in cluster 7.

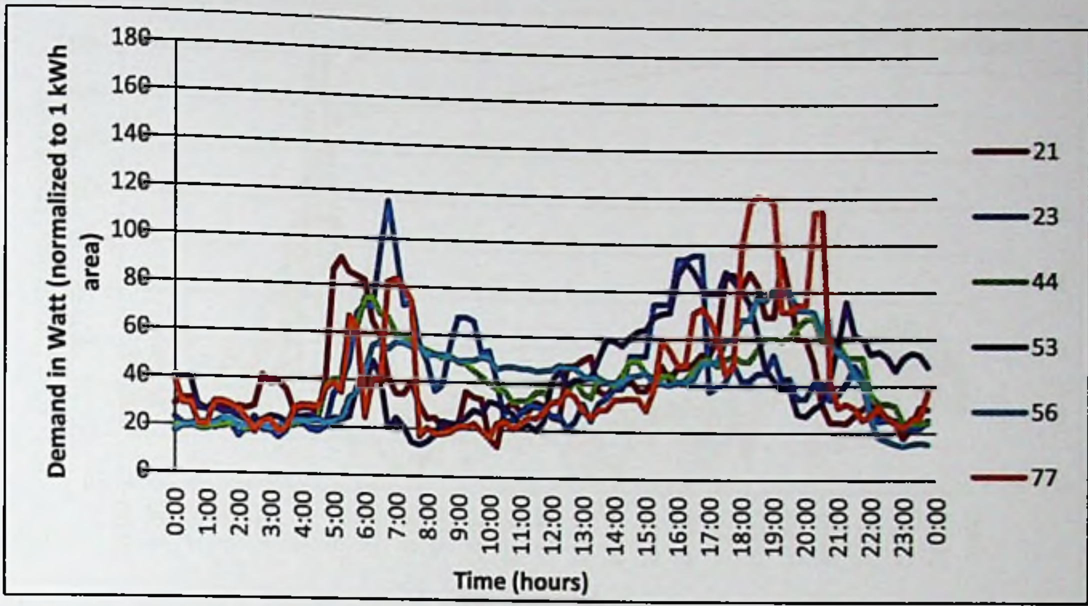


Figure 4.14 Cluster 7

As shown in figure 4.15, there is no curve specially deviated with common profile. Therefore Representative normalized curve for cluster 7 is shown in figure 4.16. This curve shows morning peak and distributed evening peak. There is a considerable load in day time with respect to rest of the time for this customer group.

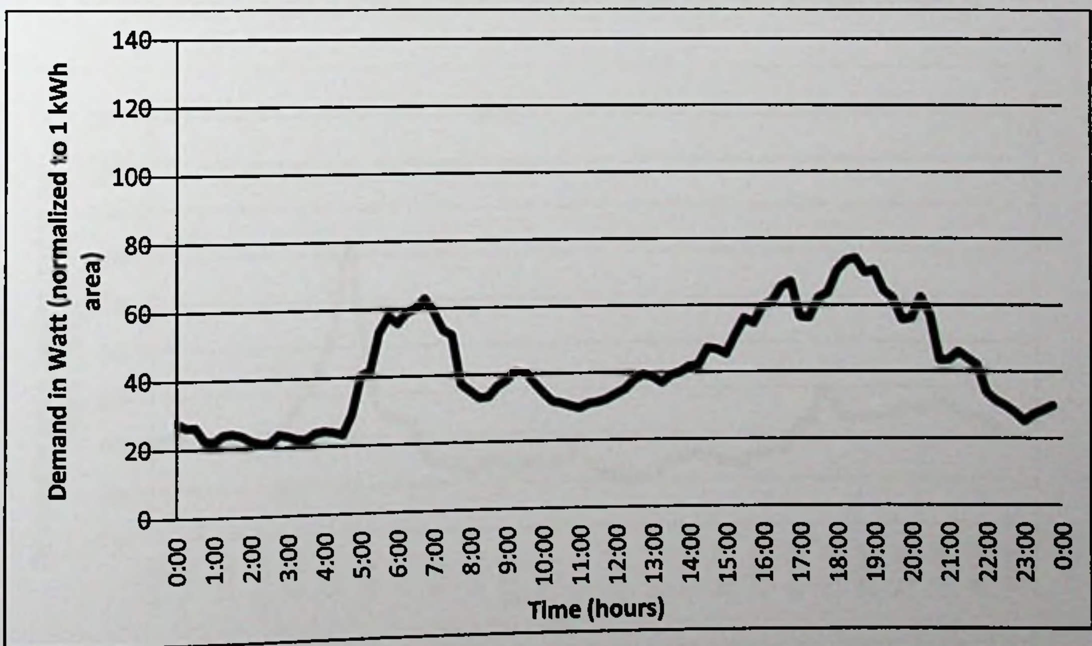


Figure 4.15 Cluster 7 Representative Curve

Figure 4.17 shows the normalized load profiles in cluster 9.

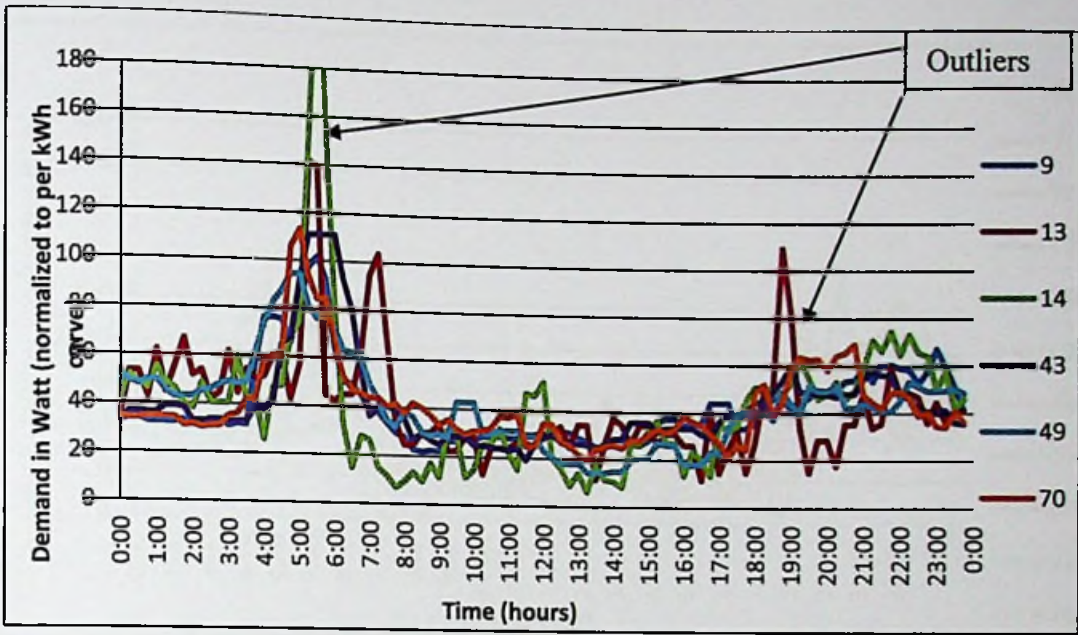


Figure 4.16 Cluster 9

As shown in figure 4.17, the curve number 13 and 14 are outliers because those are deviated from common profile of that group. Therefore those curves were not considered when obtaining representative normalized load profile for cluster 9. Representative normalized curve for cluster 9 is shown in figure 4.18. There is a huge morning peak with respect to other time periods for this customer group.

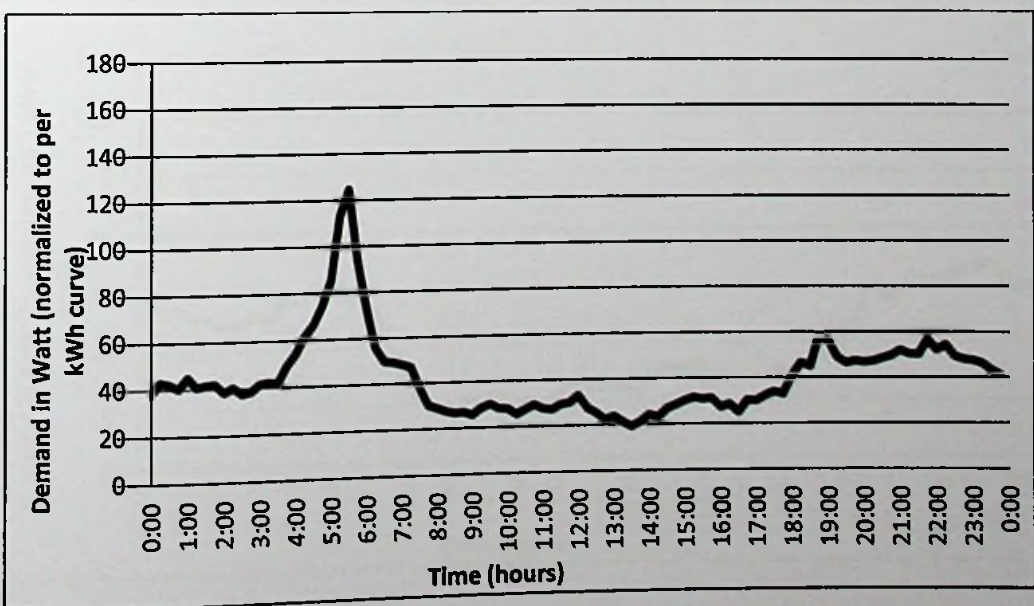


Figure 4.17 Cluster 9 Representative Curve

Figure 4.19 shows the normalized load profiles in cluster 10.

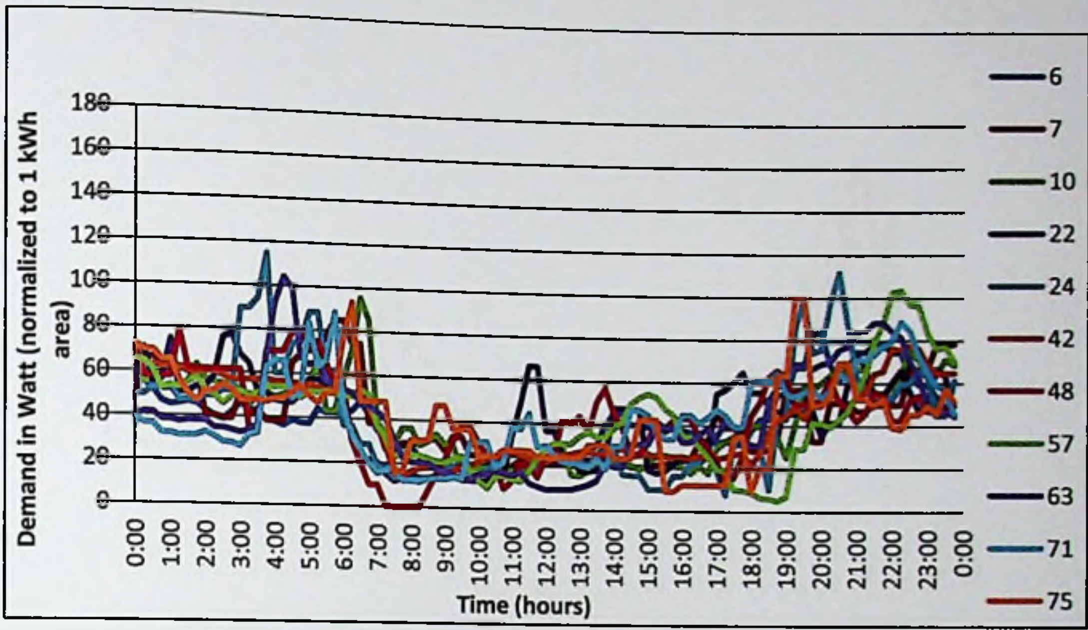


Figure 4.18 Cluster 10

As shown in figure 4.19, there is no curve specially deviated with common profile. Therefore Representative normalized curve for cluster 10 is shown in figure 4.20. There is night time energy consumption than day time loading for this customer group.

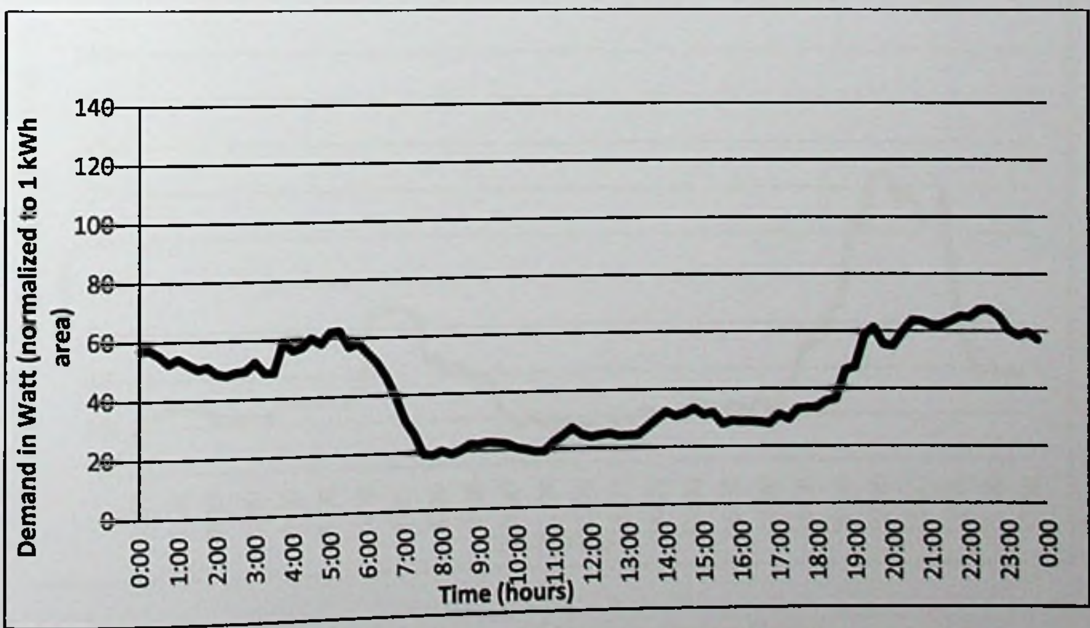


Figure 4.19 Cluster 10 Representative Curve

Figure 4.21 shows the normalized load profiles in cluster 12.

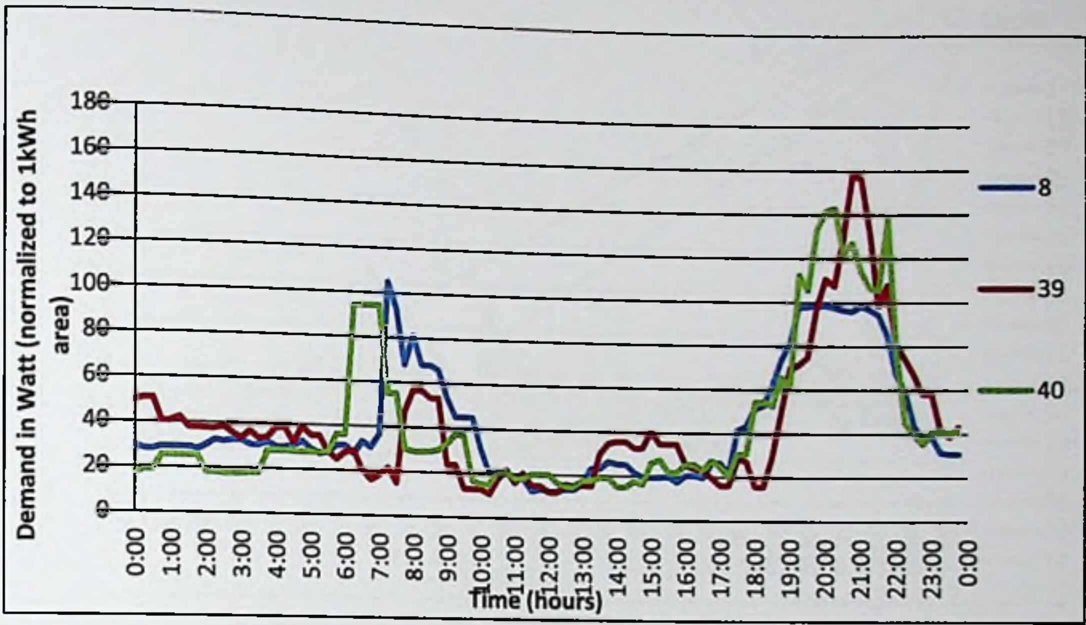


Figure 4.20 Cluster 12

As shown in figure 4.21, there is no curve specially deviated with common profile. Therefore Representative normalized curve for cluster 12 is shown in figure 4.22. This customer group shows evening peak rather than morning peak.

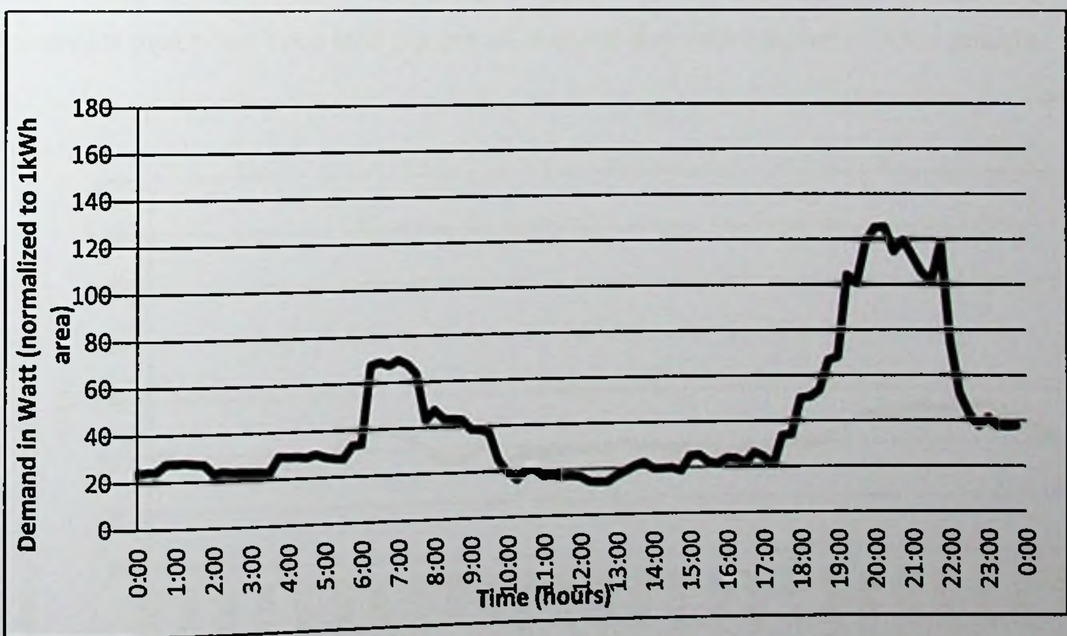


Figure 4.21 Cluster 12 Representative Curve

Figure 4.23 shows the normalized load profiles in cluster 13.

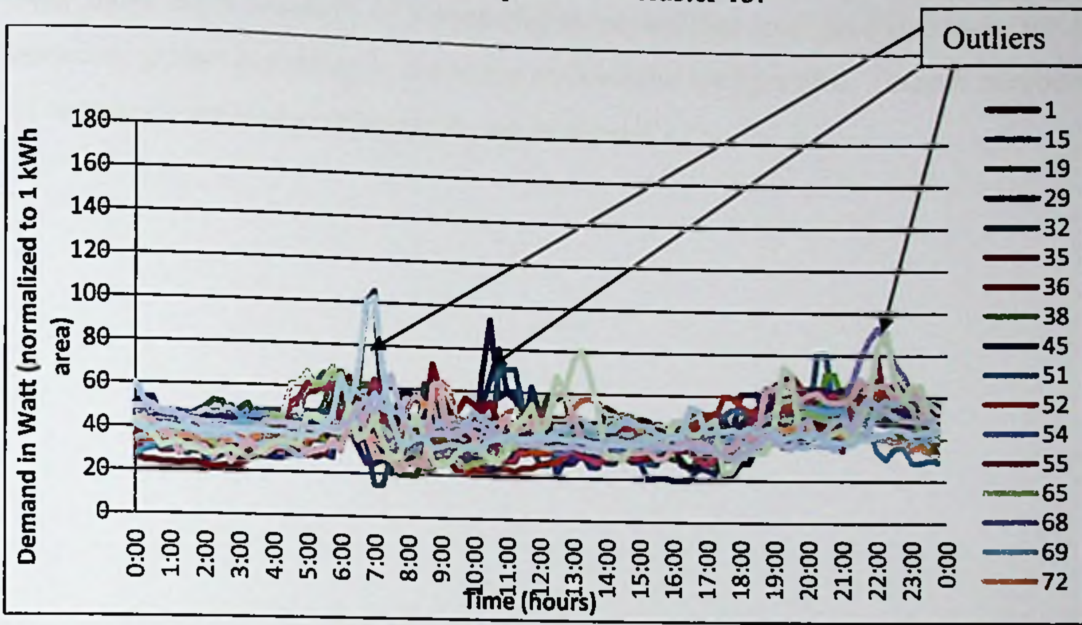


Figure 4.22 Cluster 13

As shown in figure 4.23, the curve number 15, 81, 29, 68 and 79 are outliers because those are deviated from common profile of that group. Therefore those curves were not considered when obtaining representative normalized load profile for cluster 13. Representative normalized curve for cluster 13 is shown in figure 4.24. This customer group has even load pattern all over the day with respect to other groups.

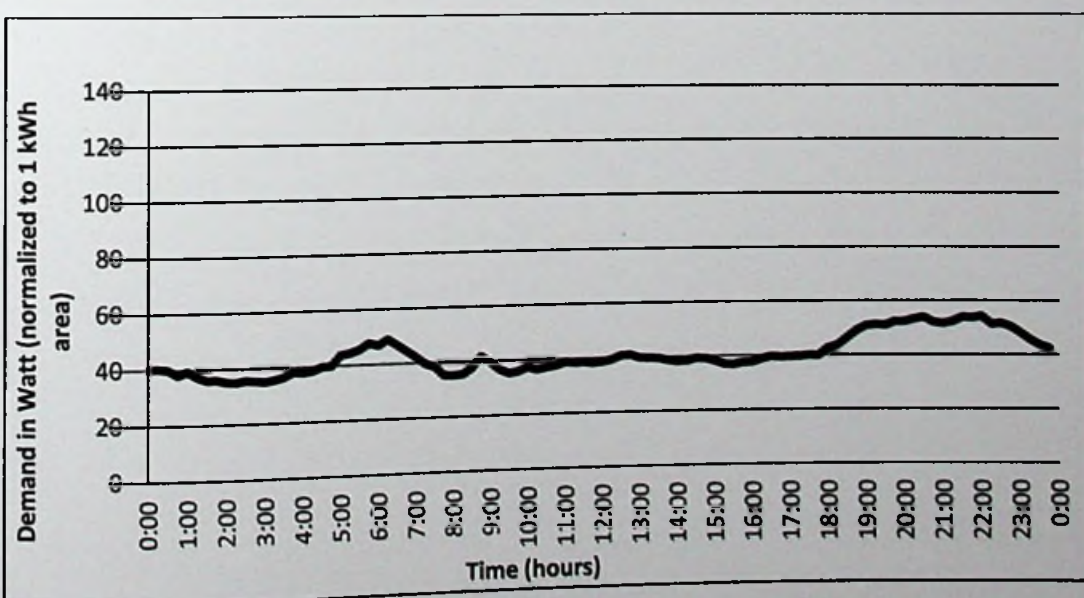


Figure 4.23 Cluster 13 Representative Curve

Now there are 8 numbers of representative normalized load profile curves for 8 customer groups according to clustering of domestic load profiles. These 8 numbers of representative load profiles are shown in appendix I.

5. DETERMINING RELEVANT GROUP OF A NEW CUSTOMER

5.1. Defining of names for clusters

The survey information was analyzed to find out similarities of customers in same group. Table 5.1 shows the information of customer family member composition details of each cluster.

Table 5.1 Family Percentage of Members

Cluster identification	Family Member composition			
	Students	Restricted time job holder	Flexible time outside job holder	At home people (self emp./retired/ unemployed)
Cluster 1	40%	90%	20%	90%
Cluster 9	100%	50%	50%	83%
Cluster 7	50%	66%	50%	100%
Cluster 12	0%	0%	100%	66%
Cluster 2	27%	45%	36%	91%
Cluster 6	25%	0%	25%	100%
Cluster 10	54%	45%	54%	63%
Cluster 13	58%	76%	30%	100%

In cluster 1, there are 40% of families which have students, 90% of families which has restricted time job holders, 20% of families which have flexible time outside job holders and 90% of families which has people at home in day time. Likewise other clusters have family percentages with members of member categories mentioned in table 5.1.

According to the statistical information in table 5.1, possibility of having students in cluster 1 is 40%, possibility of having restricted time job holders is 90%, possibility of having flexible time outside job holders is 20% and possibility of having people at home in day time is 90%. Likewise other clusters also have possibilities presented in table 5.1 of having those types of members.

Final use of this table is, if family member information is available for a new customer, the possible cluster of that new customer could be selected by referring this table. Therefore percentages values of table 5.1 should be converted to 'yes/No' statements, because the family member composition can be collected in 'Yes/No' form.

If the member composition is ≥80%, it is considered as 'Yes' and if member composition is ≤20%, it is considered as 'No'. In between 20% and 80%, is mentioned as Yes/No (may be), because the relevant cluster cannot be defined by that member information. So it is a 'don't care' information when finding relevant cluster of a new customer. Converted table to Yes/No form is shown in table 5.2.

Table 5.2 Family availability of Members

Cluster identification number	**Family Member composition**			
	Students	**Restricted time job holder**	**Flexible time outside job holder**	**At home people (self emp./retired/ unemployed)**
Cluster 1	Yes/No	Yes	No	Yes
Cluster 9	Yes	Yes/No	Yes/No	Yes
Cluster 7	Yes/No	Yes/No	Yes/No	Yes
Cluster 12	No	No	Yes	Yes/No
Cluster 2	Yes/No	Yes/No	Yes/No	Yes
Cluster 6	Yes/No	No	Yes/No	Yes
Cluster 10	Yes/No	Yes/No	Yes/No	Yes/No
Cluster 13	Yes/No	Yes/No	Yes/No	Yes

Cluster identification names for these 8 clusters are defined by considering family member information and the specialty of representative load profile pattern. First 4 clusters mentioned in table 5.3 are defined by considering family member information. Other 4 clusters are defined by special features of load representative load profile.

Table 5.3 Clsuter Identification Names

Cluster	Cluster identification number	Cluster name
Cluster 1	CL1	**Time restricted Job holders**
Cluster 9	CL9	**Students**
Cluster 7	CL7	**No job / retired people /self emp. (at home)**
Cluster 12	CL12	**Flexible time outside job holders**
Cluster 2	CL2	**Evening load consumers**
Cluster 6	CL6	**Noon load consumers**
Cluster 10	CL10	**Night time load consumers**
Cluster 13	CL13	**Even load consumers**

If a new customer to be in cluster 1, that family should have restricted time job holders. Therefore that cluster is named as **"Time restricted Job holders"**. These names does not mean that family has only the time restricted job holders, and that family may have other types of members as well according to table 5.2. These names are just for identification of clusters. The algorithm to find out relevant cluster of a new customer according to customer information is described under next topic.

If a new customer to be in cluster 9, that family should have students. Therefore that cluster is named as **"Students"**.

If a new customer to be in cluster 7, that family should have people at home in day time. There is no relevance of having or not other types of members for that cluster. Therefore that cluster is named as "**No job / retired people /self emp. (at home)"**.

If a new customer to be in cluster 12, that family should have flexible time outside job holders. Therefore that cluster is named as **"Flexible time outside job holders".**

For cluster 2, 6, 10, 13, there is no strong relationship to family member information. Therefore those clusters are named according to peaking of their power consumption according the usage of electric equipments. Figure 5.1 shows the representative load profiles of those 4 clusters. To find out relevant cluster of a new customer, the electric equipments and used time information are needed due to this reason.

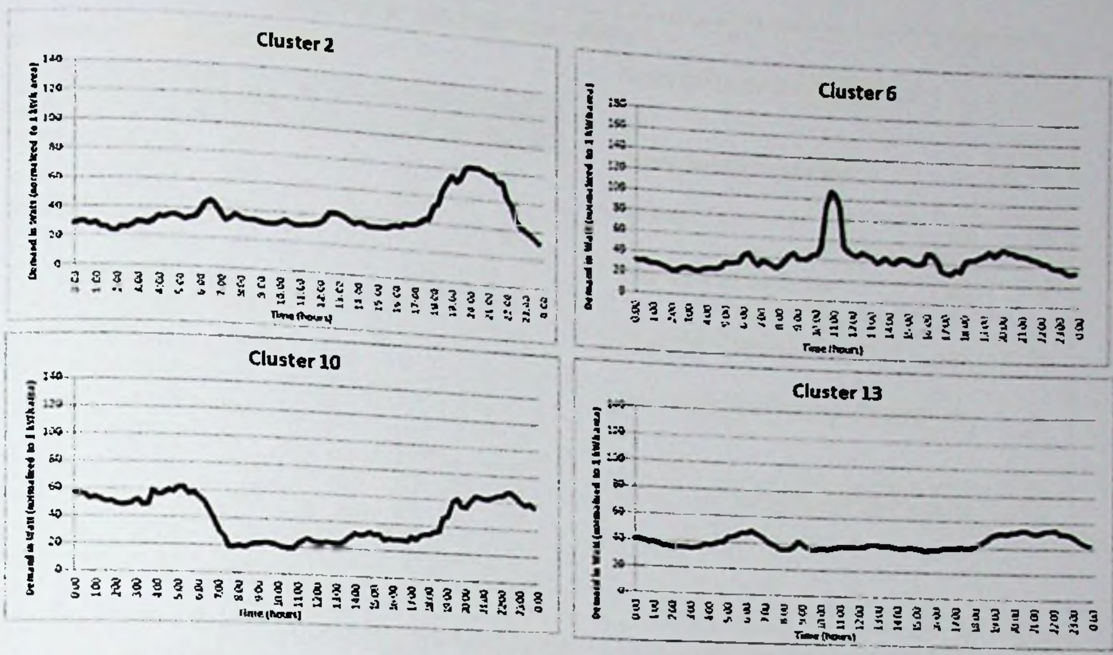


Figure 5.1 Equipment Related Clusters

Load pattern of cluster 2 has a peak demand at evening from 6.00 pm to 10.00 pm and this cluster is named as "**Evening load consumers**".

Load pattern of cluster 6 has a noon peak from 10.00 am to 12.00 pm and this cluster is named as "**Noon load consumers**".

Load pattern of cluster 10 has a higher night time power demand before 6.00am and after 6.00pm and this cluster is named as "**Night time load consumers**".

Load pattern of cluster 13 has even power demand all over the time with slightly higher demand in morning and evening. Therefore this cluster is named as "**Even load consumers**". These names and cluster ID mentioned in table 5.3 will be used in next chapters.

5.2. Determining relevant cluster according to customer information

MATLAB software application was developed to implement the algorithm of selection of representative load profile (out of 8 above mentioned) according to customer information. The graphical user interface (GUI) is shown in figure 5.2. MATLAB code is attached in appendix A.

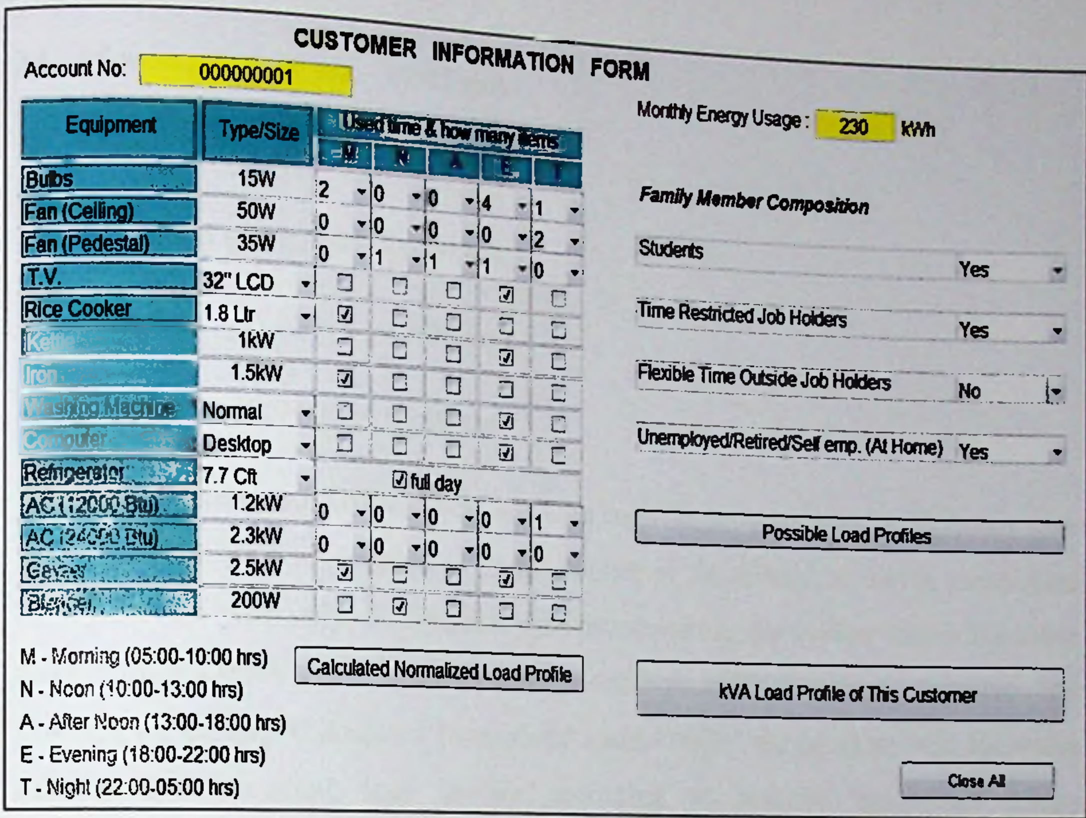


Figure 5.2 Customer Information Form

Account number of a customer is a unique number that can be referred to any database of the power distribution utility to identify the customer and save information. In this algorithm, "Monthly energy usage" is required to calculate actual load profile of a customer by representative normalized load profile curve.

In right hand side of GUI, the customer family member composition is to be input by the dropdown list. If we have only the family member composition, we cannot find the relevant customer group of a new customer exactly; there may be some of possible load profiles out of 8. Those will be appeared by clicking on button "Possible Load Profiles".

Left hand side of GUI is to input information about the electric equipments used by the customer and the usual time of using those equipments. This information form is prepared to obtain minimum set of data to collect from customers to estimate accurate and usable load profile. Collecting more data for electric equipment usage is not practical and that is not essential and valid, because customers cannot exactly answer those deep questions. The time intervals are defined as follows into 5 slots, M – Morning (05:00 am – 10:00 am)

N – Noon (10:00 am – 01:00 pm)

A – Afternoon (01:00 pm – 06:00 pm)

E - Evening (06:00 pm – 10:00 pm)

T – Night (10:00 pm – 05:00 am)

Equipments such as bulbs and fans are used in different quantities in different time intervals. The GUI is prepared to input number of items used in above mentioned time intervals separately. Equipments such as television, rice cooker and refrigerator have different sizes. There is a selection for size/type in GUI for those categories. By clicking on button "Calculated Normalized Load Profile" the program will show the calculated normalized load profile according to electric equipment usage information.

By clicking on the button "kVA Load Profile of This Customer" the program will select the nearest load profile to 'Calculated Normalized Load Profile' among 'Possible Load Profiles' by comparing Squared Euclidean Distances between each cluster centroid and calculated normalized load profile point in 96 dimensional space.

Load factors of equipments, power factors, and average operating times of different equipments are considered inside the program. As an example; operating time for rice cooker is about 30 minutes, for kettle is about 10 minutes, for Geyser is about 15 minutes, for bulbs, fans, T.V, Computer is specified time, for refrigerator is turned on for all over the time.

Block diagram for estimation of kVA load profile for a new customer with only family information and energy usage information is shown in figure 5.3.

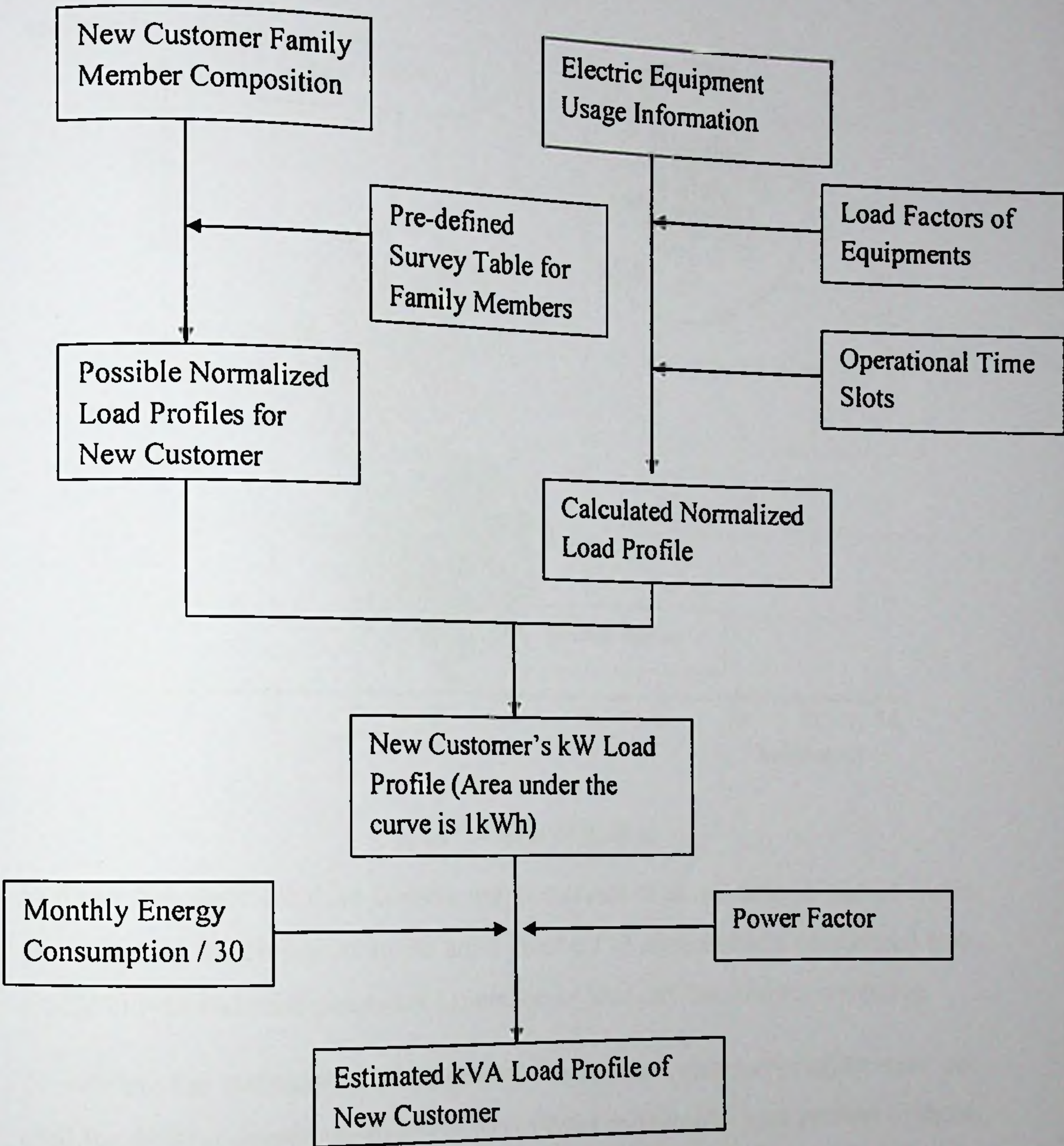


Figure 5.3 Load Profile Estimation Methodology

5.3. Method validation

To validate the results obtained by this methodology, the estimated normalized load profile curve of a new customer obtained by this methodology and actual load profile captured by the smart meter have to be compared. Let's consider an example only with two attributes (2 dimensional) for easy of graphical representation.

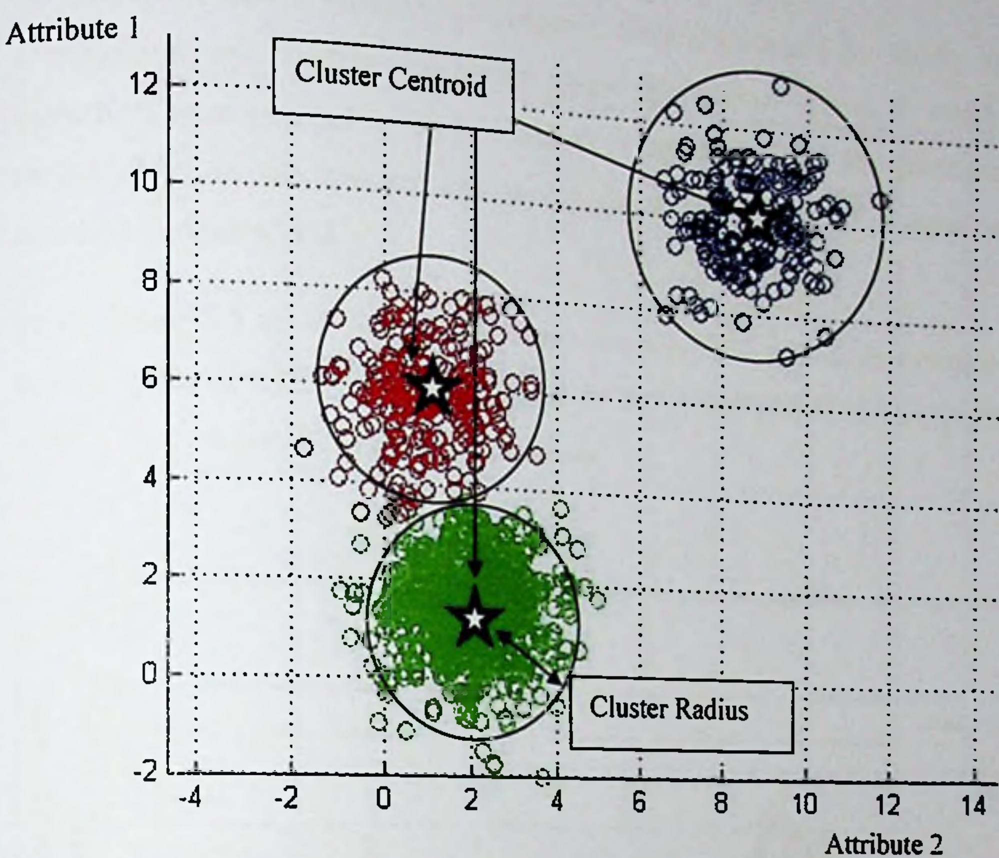


Figure 5.4 Cluster Radius

In figure 5.4, there are three clusters and 3 centroids of all clusters are named. As in this research, the cluster centroids are 8 numbers of representative normalized load profile curves and other points are 81numbers of load profiles used for clustering.

To validate the methodology, information of some new customers (which were not used for method development) has to be collected and relevant load profiles of those customers should be estimated using aforementioned methodology. Those estimated load profiles of each customer should be a cluster centroid out of 8 cluster centroids relevant to representative normalized load profile curves. Boundary of each cluster is the cluster radius which is the maximum Euclidean distance from centroid to a point belongs to cluster. If the point relevant to the actual load profile captured by smart meter is within the cluster radius of specific cluster, we can say the new customer belongs to that cluster. Hence we have selected the correct cluster for a new customer out of 8 clusters by the developed methodology.

As examples, new 8 numbers of customers were interviewed for family member composition and electric equipment usage information. Those 8 numbers of customers were external to the 81 numbers of customers used for development of methodology in this research and survey sheets information of 8 customers are attached in Appendix B.

From figure 5.5 up to figure 5.12 shows the comparison between estimated load profile by the developed methodology and the actual load profile curve captured by a smart meter for new 8 numbers of customers.

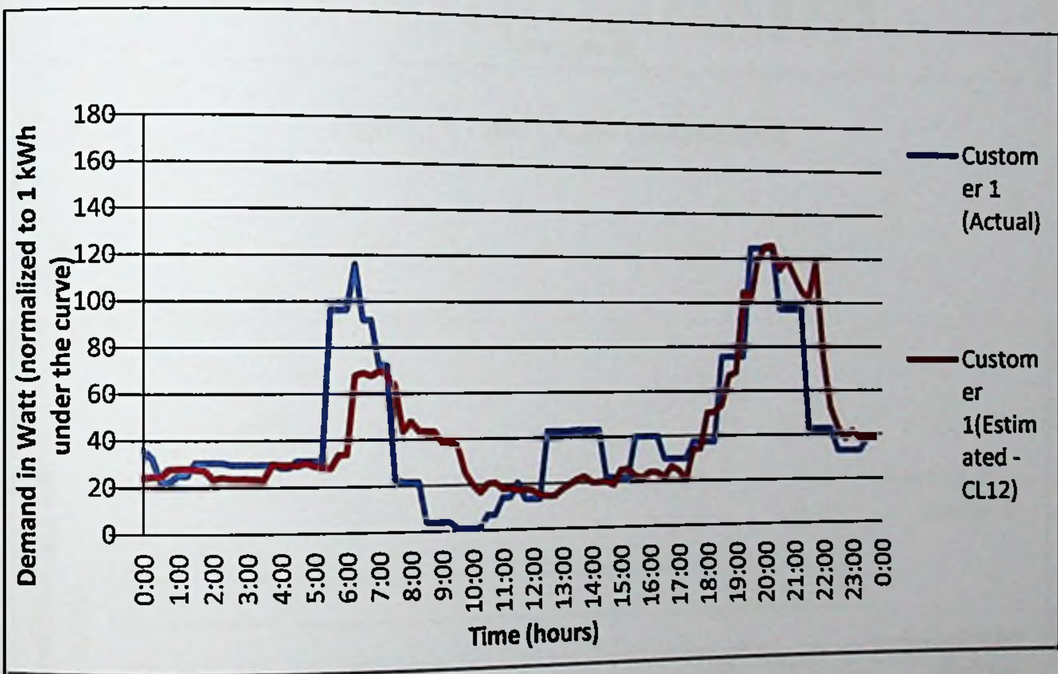


Figure 5.5 Load Profile Comparison 1

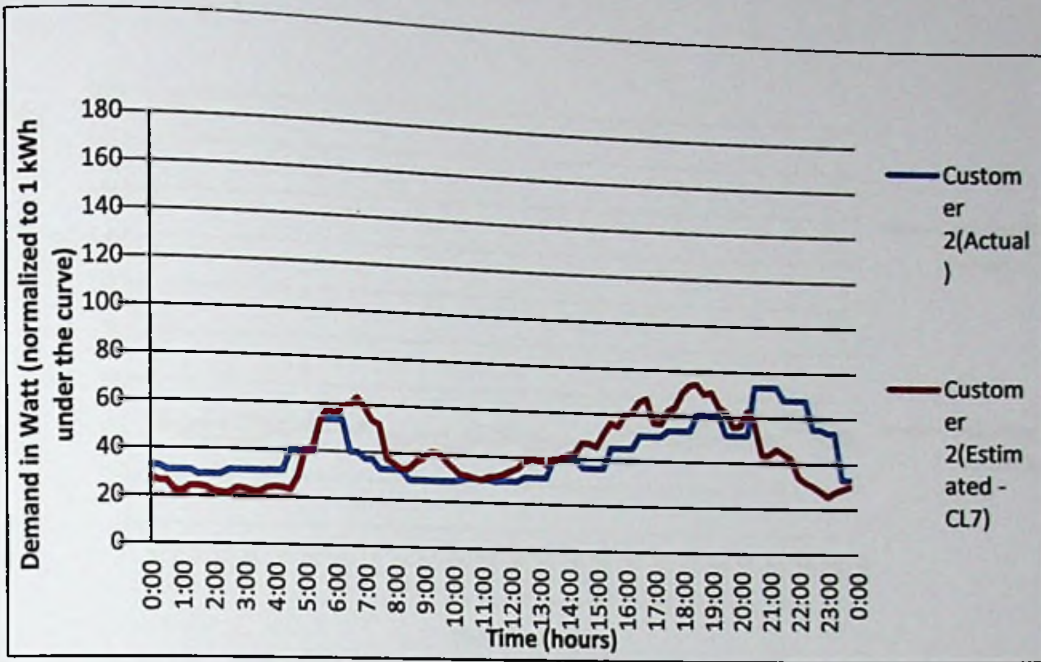


Figure 5.6 Load Profile Comparison 2

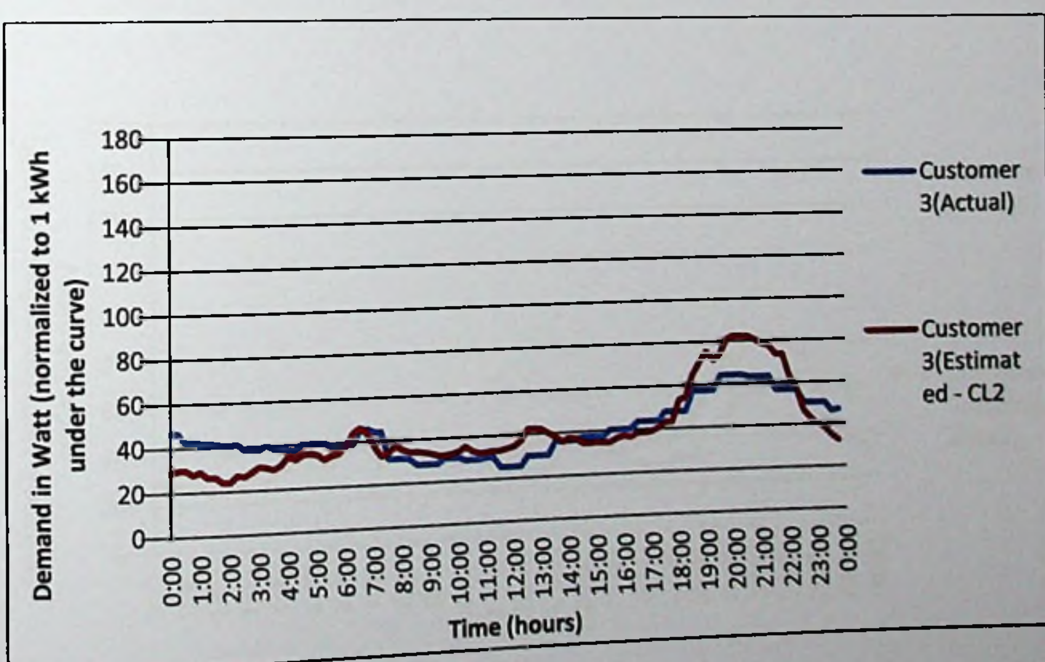


Figure 5.7 Load Profile Comparison 3

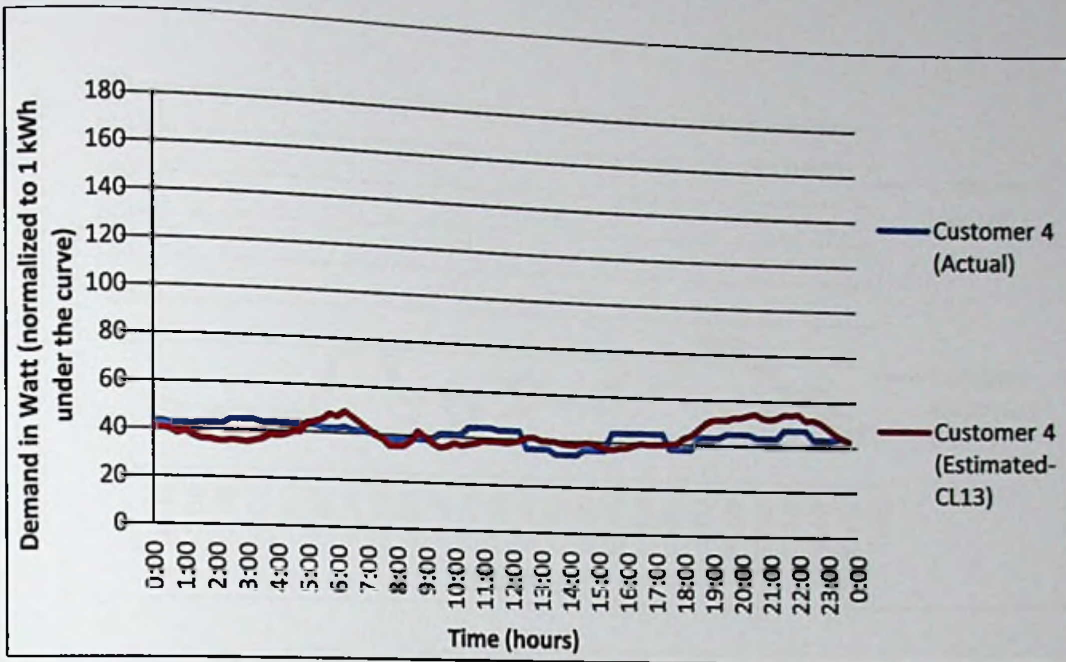


Figure 5.8 Load Profile Comparison 4

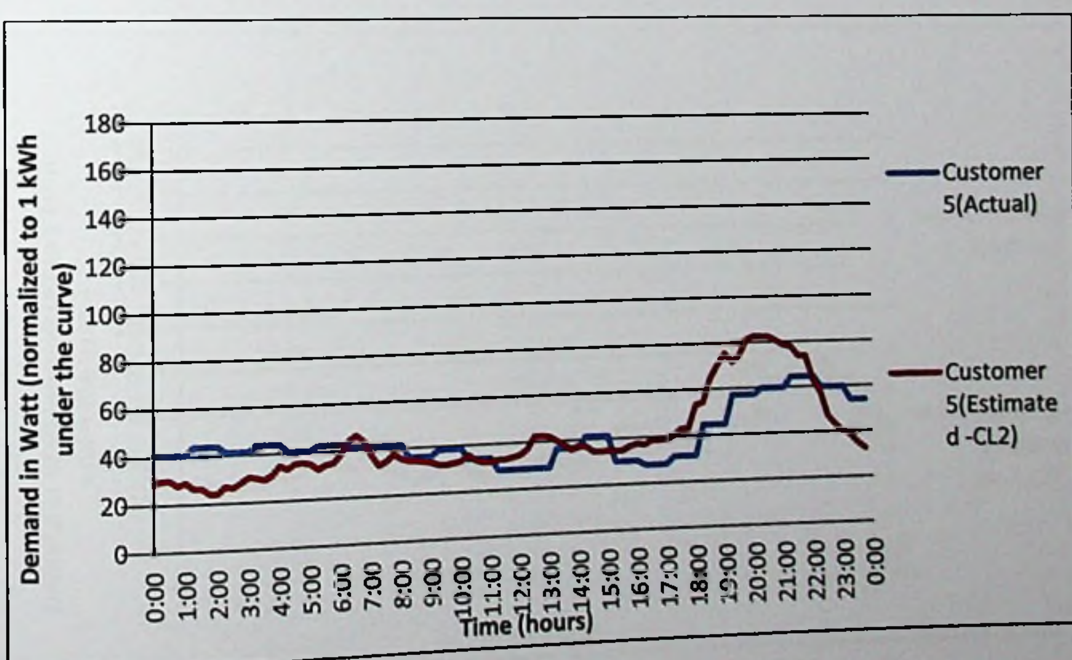


Figure 5.9 Load Profile Comparison 5

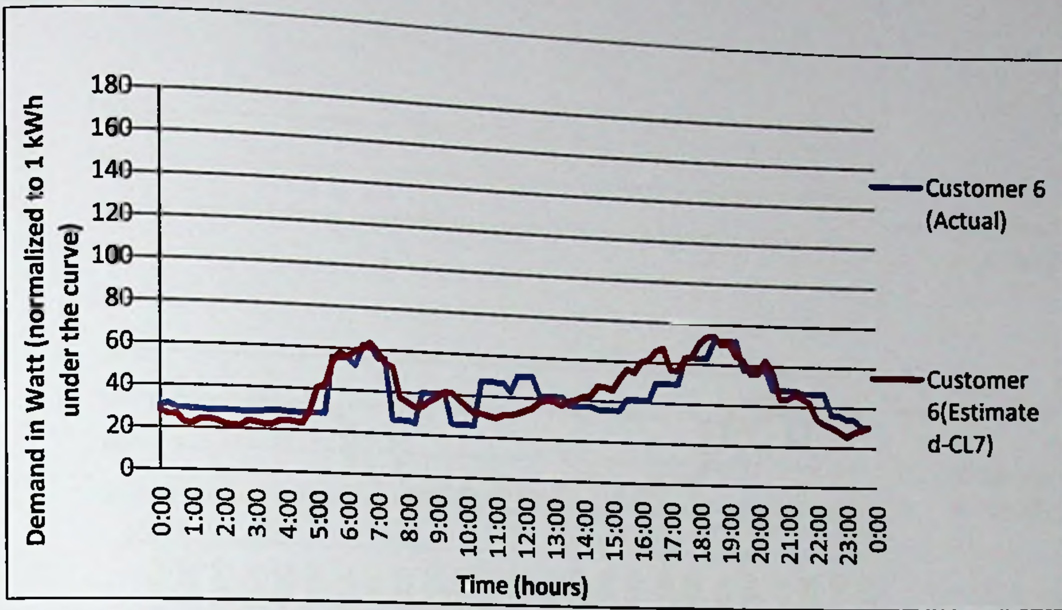


Figure 5.10 Load Profile Comparison 6

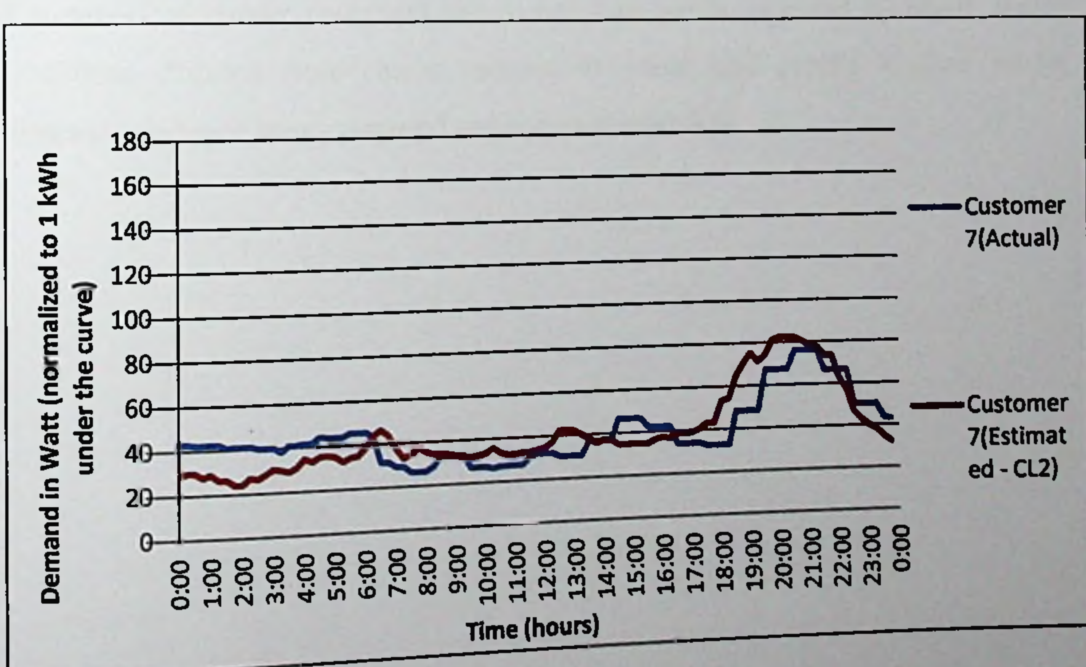


Figure 5.11 Load Profile Comparison 7

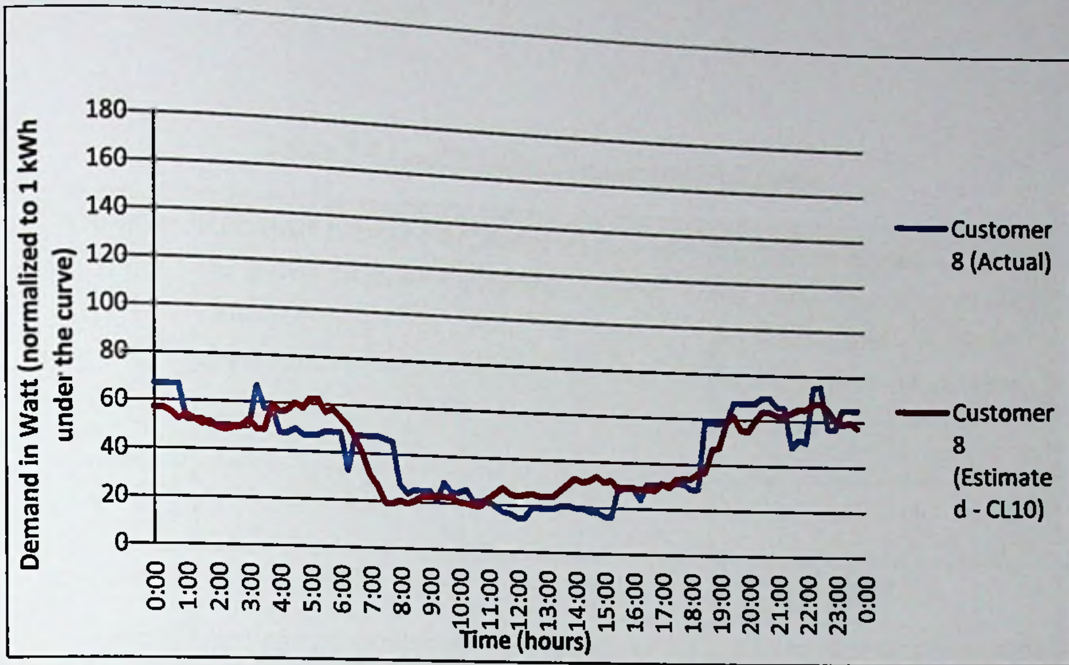


Figure 5.12 Load Profile Comparison 8

To obtain a numerical value for the deviation between estimated load profile (one of 8 numbers of cluster centroids) and Actual load profile captured by smart meter, Euclidean distance from cluster centroid to actual load profile location in 96 dimensional space were compared and shown in table 5.4.

Table 5.4 Load Profile Comparison Validation

Customer No:	Relevant cluster ID for new customer deduced by algorithm	Cluster radius of relevant cluster	Euclidean distance between cluster centroid to actual load profile location in 96 dimensional space
Customer 1	CL 12	184	220.7112
Customer 2	CL 7	170	130.2372
Customer 3	CL 2	145	94.77481
Customer 4	CL 13	147	53.56557
Customer 5	CL 2	145	119.9091
Customer 6	CL 7	170	89.32261
Customer 7	CL 2	145	103.1271
Customer 8	CL 10	165	94.20077

According to comparison in table 5.4, all new customers' load profile points are within the cluster radius of each relevant cluster except for customer 1. As described in chapter 4, initially there were 13 clusters but only 8 clusters were selected as rich-clusters. Therefore these types of exceptions may happen when estimating the load profile, because 5 numbers of poor-clusters were omitted.

But this methodology selects the closest cluster centroid when selecting relevant cluster for new customer. Though above customer 1 is outside the cluster radius, the closest cluster for customer 1 is selected as Cluster 12. Therefore this methodology selects a fare result for any customer. In comparison 1, figure 5.5, we can see that, this methodology selected a most suitable load profile for customer 1.

6. LV FEEDER POWER LOSS CALCULATION

In this thesis, chapter 1 to chapter 5 were discussed the importance of estimating domestic load profiles. One application of estimated load profile is presented in this chapter. Calculation of low voltage feeder power loss is very important for electricity distribution utility. Inaccuracy of feeder line loss calculation by existing energy meters is discussed in problem statement in chapter 1.

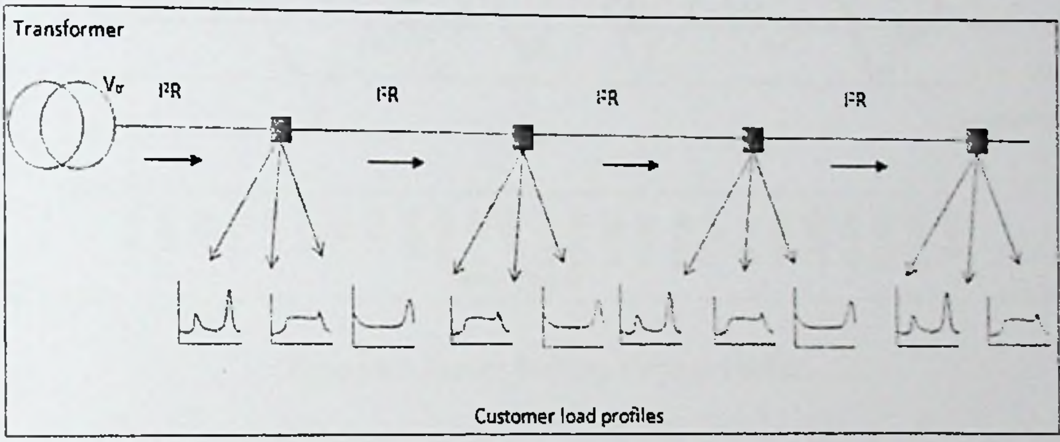


Figure 6.1 LV Feeder Power Loss

As shown in figure 6.1, the power loss of a specific feeder can be calculated using kVA load profiles of connected customers, conductor resistances of feeder sections and starting point voltage profile. Customer load profiles can be estimated according to the methodology presented in this thesis. Conductor resistances can be calculated using low voltage feeder details in geographical information system of distribution utility. Starting point voltage profile is captured by substation meter. Forward-Backward Sweep Method can be applied for calculation of currents of each feeder section and total line power loss. To obtain a curve for feeder power loss in 24 hour domain, this calculation should be repeated in 15 minutes intervals, since the load profiles of customer and voltage profile transformer are in 15 minutes intervals.

As an example, small feeder section was selected in Pita-Kotte area where all 8 numbers of customers connected to feeder section have smart meters with load profile recording facility. Additionally, a substation meter was installed to capture voltage profile at the starting point of that feeder section. Voltage variation of feeder starting point is shown in figure 6.2 which is captured by installed substation meter. Set of voltage data in 24 hour domain is attached in Appendix F.

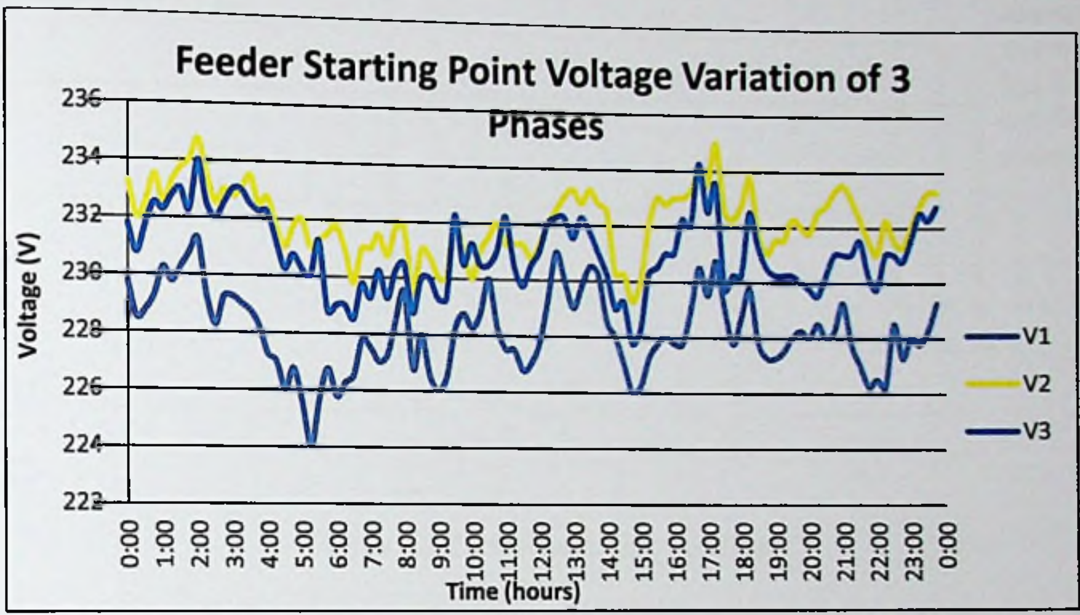


Figure 6.2 Feeder Starting Voltage Profile

Figure 6.3 shows the line diagram of selected feeder section.

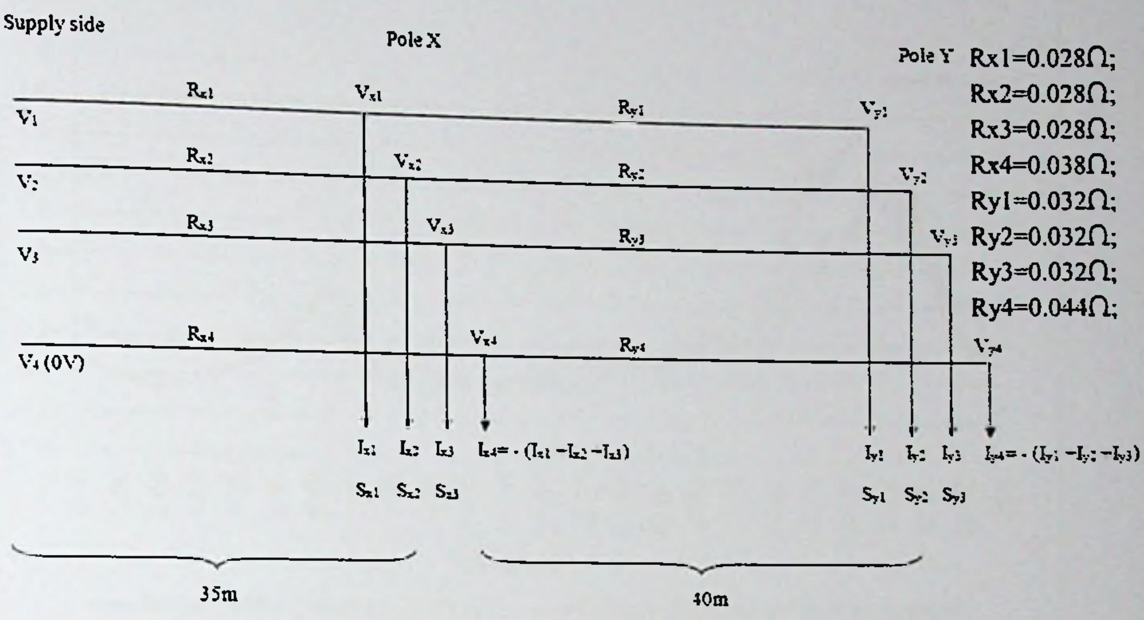


Figure 6.3 Three Phase Line Loading

Most of the domestic customers have single phase connections; therefore connected phase of each customer is important when calculating accurate figure for feeder section power loss. Current of neutral conductor is also considered in this calculation.

In initial backward sweep, voltage at pole x and pole y are assumed to be 230V. The currents of each feeder section are calculated using estimated loads of each customer. Power factor is assumed to be 0.9 when obtaining kVA load profile. Then using voltage at feeder starting point, the voltages at pole x and pole y are calculated in forward sweep. This was done for 10 iterations and currents of each section are deduced. This procedure is repeated for 96 times to obtain feeder loss curve for 24 hour domain in 15 minutes interval. MATLAB program coding for this activity is attached in Appendix E.

Feeder section power loss calculation was done by considering estimated load profile curves according to the customer information as well as actual load profiles captured by smart meters connected to those customers. The comparison of power loss profiles in two scenarios is shown in figure 6.4.

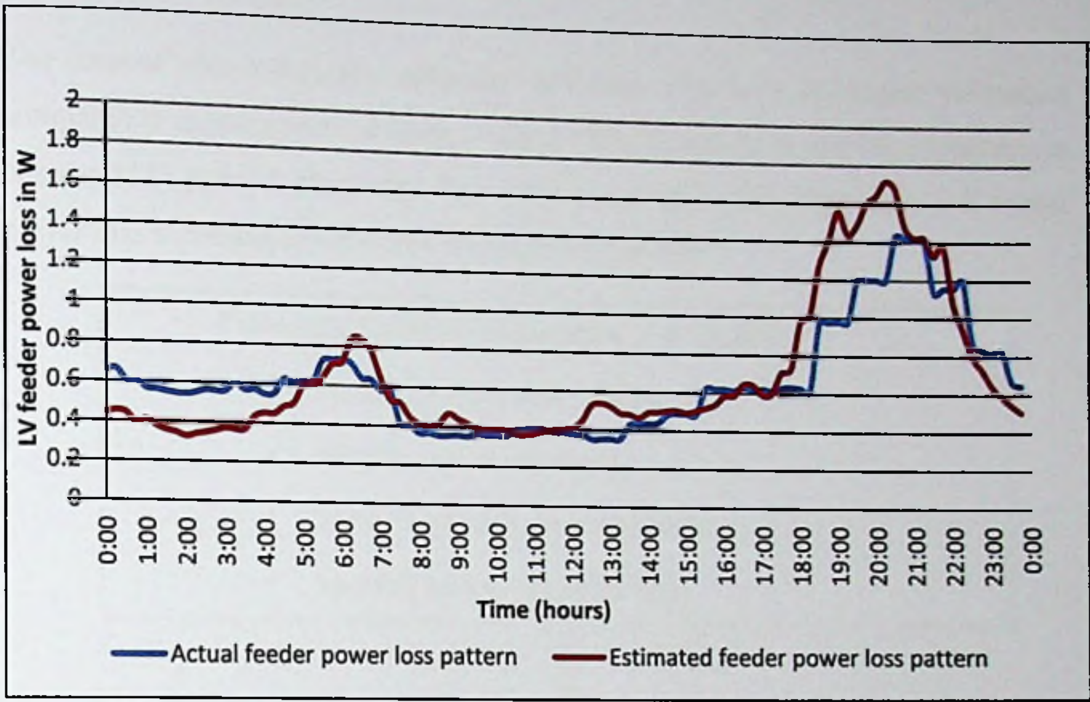


Figure 6.4 24 Hour LV Feeder Section Power Loss

The area under the curve is total energy loss of the feeder section for a 1 day. Since this is a small feeder section with only 8 customers, the figures are small in magnitude. But for a 160 kVA distribution transformer, there are about 250 numbers of customers and feeder power losses will be in considerable magnitudes.

- Total estimated energy loss of feeder section for 1 day = 15.95 Wh
- Total estimated energy loss of feeder section for 1 day = 15.42 Wh

The error of total energy loss estimation is 3.44%. This is an acceptable estimation compared to existing method (error is about 50%) described in problem statement in chapter 1. Figure 6.5 shows the flow chart to calculate time dependant LV feeder power loss according to estimated customer load profiles.

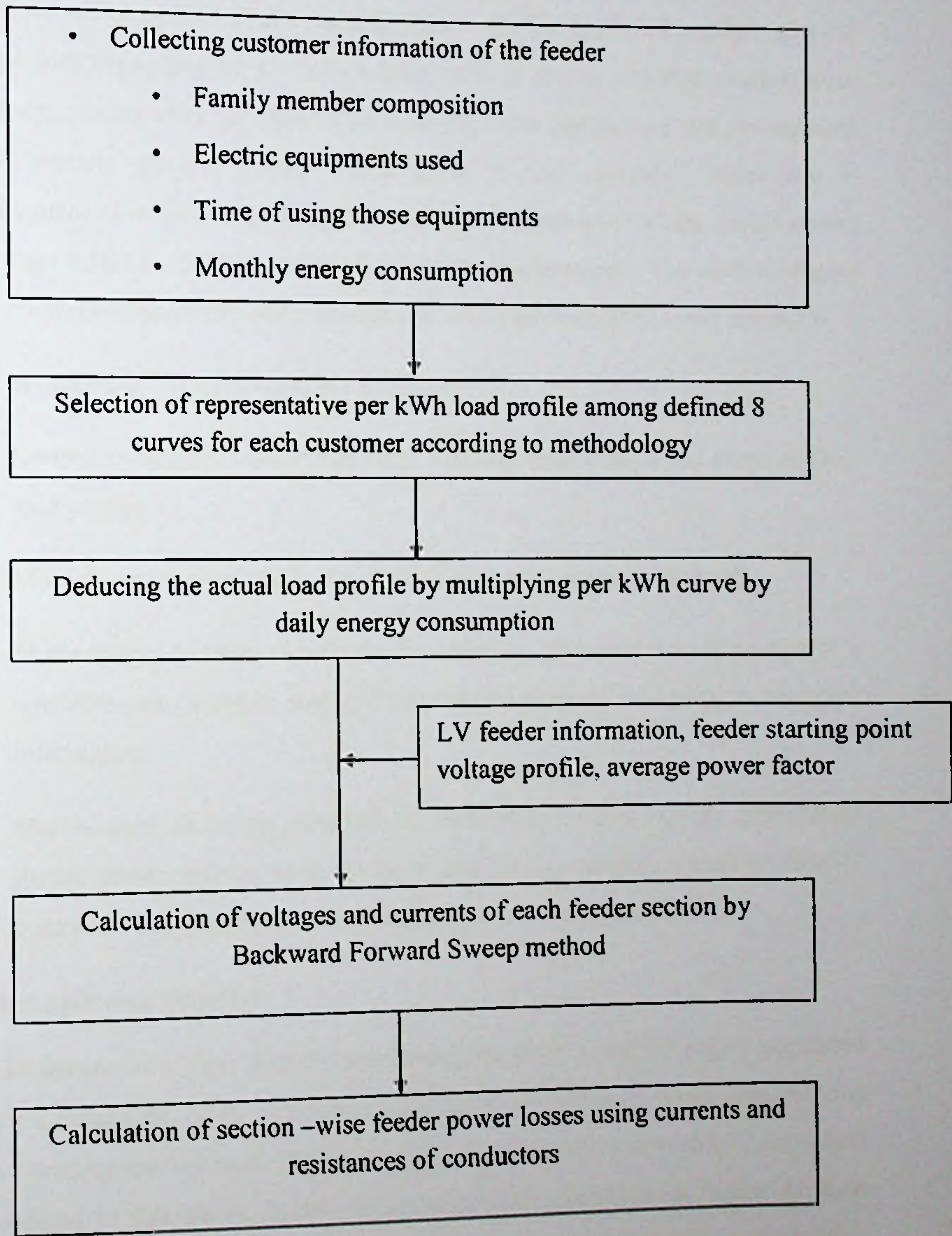


Figure 6.5 Feeder Power Loss Estimation

7. CONCLUSION

7.1. Achievement of objective and research outcome

There were two objectives of this research, one is "Defining of load profile curve for any customer depending on his information such as family member composition, electric equipments used, the time/nature of using those equipments and system data of past average monthly energy consumption of that customer" other one is "Development of methodology to estimate time dependant low voltage feeder power loss (power losses in conductors) for distribution transformers". The methodologies to achieve those objectives were proposed and validated in chapter 5 and chapter 6.

Research outcomes can be pointed out as follows

- Domestic customers are divided into 8 groups according to the shape of their load profile.
- Most representative load profile pattern for each group was defined.
- Methodology is being developed for selection of relevant load pattern of a new domestic customer among 8 pre-defined patterns, according to customer information.
- Methodology is being proposed for estimation of low voltage distribution feeder power loss curve in 24 hour domain, by estimated load profiles of customers and geographical and feeder network information.

7.2. Applicable situations

The selected area to collect data and conducting the study is neither highly populated area nor a rural area. Kotte area has a wide range of customers which has monthly energy consumption between 50kWh to 500kWh. Therefore every type of customers is considered in this study. In Sri Lanka, electricity customers are using common household equipments and those are considered when implementing methodology. Therefore this is applicable for almost every type of domestic customers.

7.3. Limitations

Load growth of a domestic customer can be considered as two scenarios. One is load deepening and other one is load widening. Load deepening means, one customer is using same electric equipments, but usage of those equipments is increased. In this scenario, the load profile pattern is not affected, but only the magnitude is increased. Load widening means, if a customer start to use new electric equipment, the load pattern is affected. Therefore the customer in the areas where load widening is dominant, those customers are moving among the different clusters. Frequent customer classification is needed for those areas.

Effect of weather is not much affected for Sri Lankan situation because there is no differentiated season change. But in rainy days, usage of air conditioners and fans is reduced with respect to normal use. This situation will affect the estimated load profiles. Another study is needed to analyze this situation.

In special cases deviated from normal household nature such as boarding complexes, we cannot consider those as domestic customers. If someone has net metering connections, electric car, then smart meters should be installed to capture actual load profiles of those customers to estimate load profiles for feeder power loss calculation.

In holydays and festival seasons, we cannot predict the behavior of the customers. The estimated load pattern is affected due to special uses of electric appliances. This is also a limitation for application of this methodology. Another study is needed to analyze these situations and load profile data is needed for several years on those special days.

7.4. Recommendations for power distribution utility

In Sri Lanka, there are nearly 5 million domestic customers. It is recommended for an electricity distribution utility to keep about 5000 smart metering customer population in a selected area to implement online load profile classification program. Information of those 5000 customers should be updated at least twice a year. Changes of load profiles of those customers can be captured once a month. If a customer changes the pattern of the load profile and moves to another customer group, utility can investigate reasons and can be added modification to the load profile defining algorithm.

Updated methodology can be applied to define unknown load profiles for rest of the customers which has not installed smart meters as described in this thesis.

References

1. Ke Wang, Mingming Zhang, Zhimin Wang, Ran Li, Furong Li, Hao Wuc, "Time of Use Tariff Design for Domestic Customers from Flat Rate by Model-based Clustering", The 6th International Conference on Applied Energy – ICAE2014

2. In Hyeob Yu, Jin Ki Lee, Jong Min Ko, Sun Ic Kim, "A method for Classification of Electricity Demands using load Profile Data". Korea Electric Power Research institute 2015

3. Gheorghe Grigoras, Ovidiu Ivanov, and Mihai Gavrilas, "Customer Classification and Load Profiling using Data from Smart Meters" 12th symposium on Neural network applications in electrical engineering, School of Engineering, University of Belgrode, Serbia, 2014

4. Christian Beckel, Leyna Sadamori, Silvia Santini, "Towards Automatic Classification of Private Households Using Electricity Consumption Data", Institute for Pervasive Computing, ETH Zurich, 8092 Zurich, Switzerland 2012

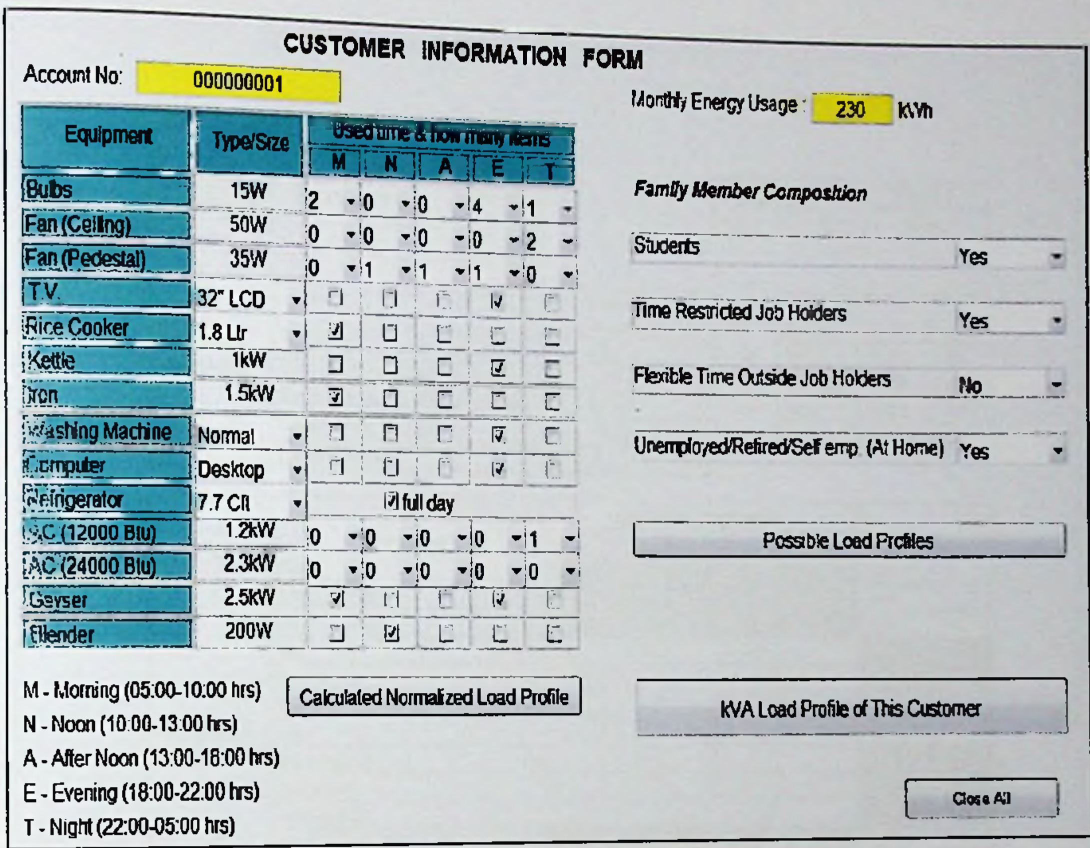
5. Fintan McLoughlin, "Characterizing Domestic Electricity Demand for Customer Load Profile Segmentation", Dublin Institute of Technology ARROW@DIT, 2013

6. Antti Mutanen, Pertti Järventausta, Matti Kärenlampi, Pentti Juuti, "Improving Distribution Network Analysis With New AMR Based Load Profiles", 22nd International Conference on Electricity Distribution Stockholm, 10-13 June 2013

7. Antti Mutanen, Maija Ruska, Sami Repo, and Pertti Järventausta, "Customer Classification and Load Profiling Method for Distribution Systems", TPWRD-00616-2010

8. T. Soni Madhulatha, "AN OVERVIEW ON CLUSTERING METHODS", IOSR Journal of Engineering 2012, Vol. 2(4) pp: 719-725.

APPENDIX A : Graphical User Interface



Program code of above interface and Matlab files are burnt in CD as Appendix A

APPENDIX B : Customer information forms of validated customers

Survey Sheet

Account No:	Customer 1 —
Transformer:	
Serial No:	

Units:	40 Units
Bill :	200 Rs

		How many members
Member composition	**Students**	No
	Restricted time job holder	No
	Flexible time out side job holder	Yes
	Unemployed people	Yes
	Retired people	No

M	Morn.	05:00 - 10:00
N	Noon	10:00 - 13:00
A	Af.Noon	13:00 - 18:00
E	Eve.	18:00 - 22:00
T	Night	22:00 - 05:00

	Equipment	Used time/ How many M	N	A	E	T	Size/ Type
Equipment usage	Bulbs	2	—	—	2	—	N/A
	Fan (Ceiling)	[illegible]	[illegible]	[illegible]	—	1	N/A
	Fan (Pedestal)					[illegible]	N/A
	TV				✓		[illegible] inch
	Rice cooker	✓					[illegible]
	Kettle						N/A
	Iron						N/A
	washing mc.						kg
	Computer				✓		Desk Tp/Lap
	Refrigerator						[illegible] feet
	AC (12000 Btu)						N/A
	AC (24000 Btu)						N/A
	Geyser						N/A
	blender						N/A

Dwelling	**Rooms**	2
	Stories	1
	Size	M
	Lighting level	good
	Closed / Open	open

Special needs:

—

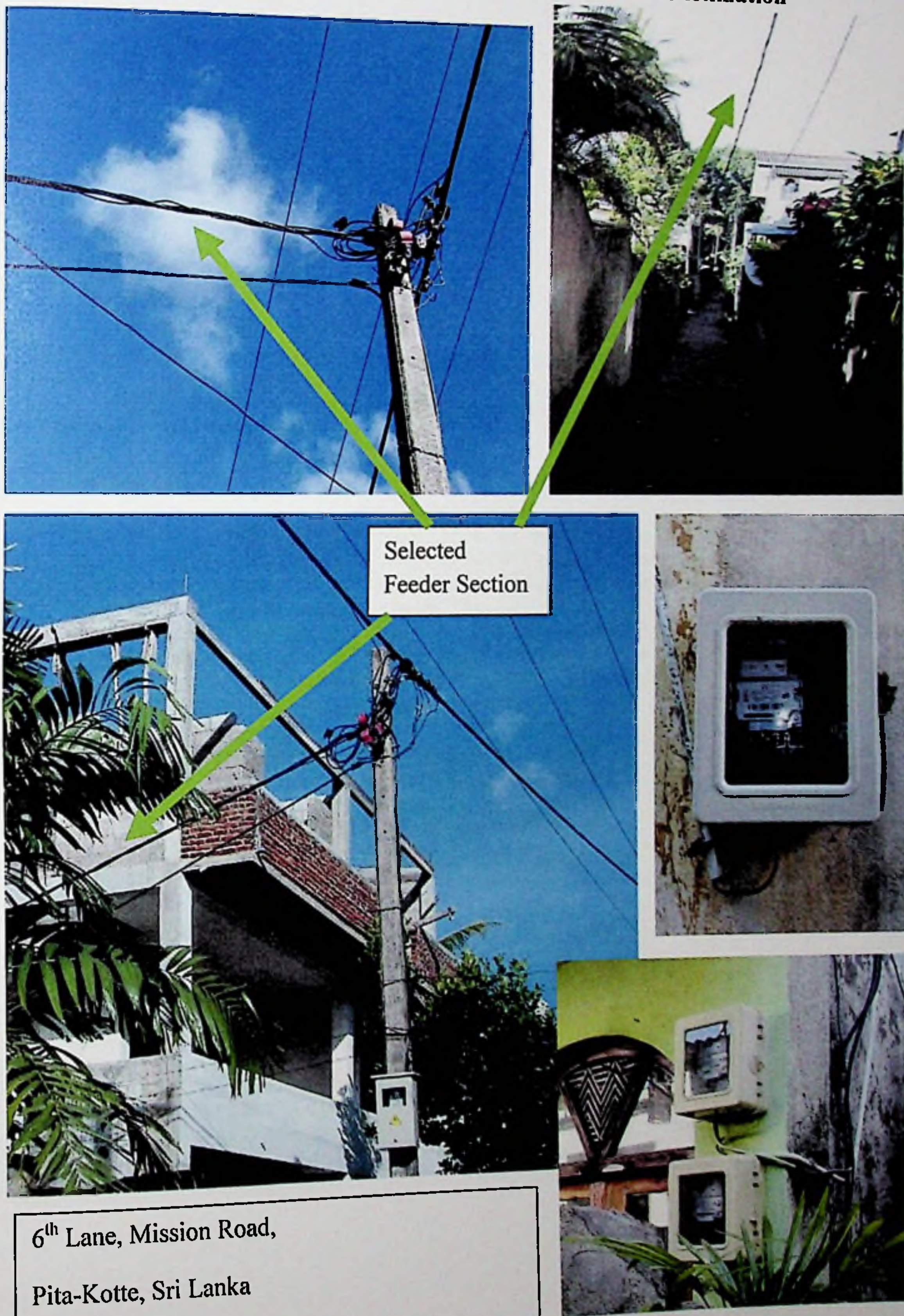
Rest of scanned copies are burnt in CD as appendix B

Appendix C: Smart meter data of Validated 8 Customers

Customer No:	1	2	3	4	5	6	7	8
Cluster	12	7	2	13	2	7	2	10
Pole No:	1	2	2	2	2	1	1	1
Phase	1	3	3	1	2	2	3	1
kWh/Day	1.076	3.161	4.519	7.136	3.775	2.863	4.977	2.770
0:00	0.039	0.104	0.211	0.303	0.153	0.086	0.213	0.185
0:15	0.035	0.103	0.210	0.308	0.153	0.090	0.213	0.185
0:30	0.024	0.098	0.191	0.303	0.153	0.084	0.210	0.185
0:45	0.024	0.098	0.191	0.303	0.153	0.084	0.210	0.185
1:00	0.027	0.098	0.190	0.301	0.153	0.083	0.211	0.145
1:15	0.027	0.098	0.190	0.301	0.165	0.083	0.211	0.145
1:30	0.033	0.092	0.187	0.304	0.165	0.081	0.202	0.145
1:45	0.033	0.093	0.184	0.304	0.165	0.082	0.203	0.138
2:00	0.033	0.093	0.184	0.304	0.165	0.082	0.203	0.138
2:15	0.033	0.093	0.184	0.304	0.155	0.082	0.203	0.138
2:30	0.032	0.098	0.173	0.318	0.155	0.081	0.199	0.138
2:45	0.032	0.098	0.173	0.318	0.155	0.081	0.199	0.138
3:00	0.032	0.098	0.173	0.318	0.155	0.081	0.199	0.138
3:15	0.032	0.098	0.181	0.316	0.165	0.082	0.190	0.185
3:30	0.032	0.098	0.173	0.309	0.165	0.083	0.203	0.158
3:45	0.032	0.098	0.173	0.309	0.165	0.083	0.203	0.158
4:00	0.030	0.099	0.170	0.309	0.165	0.082	0.206	0.132
4:15	0.030	0.099	0.170	0.309	0.153	0.082	0.206	0.132
4:30	0.033	0.127	0.181	0.307	0.153	0.081	0.220	0.137
4:45	0.033	0.125	0.181	0.307	0.153	0.082	0.217	0.130
5:00	0.033	0.125	0.181	0.307	0.153	0.082	0.217	0.130
5:15	0.033	0.125	0.181	0.307	0.160	0.082	0.217	0.130
5:30	0.103	0.169	0.174	0.299	0.160	0.159	0.226	0.134
5:45	0.103	0.169	0.174	0.299	0.160	0.159	0.226	0.134
6:00	0.103	0.169	0.174	0.299	0.160	0.159	0.226	0.134
6:15	0.125	0.169	0.175	0.303	0.155	0.148	0.214	0.090
6:30	0.099	0.126	0.205	0.294	0.155	0.176	0.156	0.131
6:45	0.099	0.126	0.205	0.294	0.155	0.176	0.156	0.131
7:00	0.077	0.117	0.198	0.289	0.155	0.159	0.144	0.131
7:15	0.077	0.117	0.198	0.289	0.155	0.159	0.144	0.131
7:30	0.024	0.104	0.141	0.270	0.155	0.076	0.130	0.128
7:45	0.023	0.104	0.141	0.270	0.155	0.076	0.130	0.125
8:00	0.023	0.104	0.141	0.270	0.155	0.076	0.130	0.075
8:15	0.023	0.103	0.140	0.269	0.135	0.072	0.141	0.065
8:30	0.004	0.091	0.127	0.274	0.135	0.115	0.164	0.069
8:45	0.004	0.091	0.127	0.274	0.135	0.115	0.164	0.069
9:00	0.004	0.091	0.127	0.274	0.135	0.115	0.164	0.069
9:15	0.004	0.091	0.127	0.274	0.145	0.115	0.164	0.057
9:30	0.001	0.091	0.138	0.294	0.145	0.073	0.134	0.079
9:45	0.001	0.091	0.138	0.294	0.145	0.073	0.134	0.068
10:00	0.001	0.091	0.138	0.294	0.145	0.073	0.134	0.068
10:15	0.001	0.092	0.131	0.294	0.127	0.072	0.130	0.072
10:30	0.007	0.095	0.133	0.315	0.127	0.135	0.134	0.059
10:45	0.007	0.095	0.133	0.315	0.127	0.135	0.134	0.059
11:00	0.015	0.094	0.136	0.316	0.127	0.134	0.135	0.060
11:15	0.015	0.094	0.136	0.316	0.108	0.134	0.135	0.053

11:30	0.022	0.092	0.114	0.310	0.108	0.122	0.151	0.047
11:45	0.014	0.093	0.114	0.310	0.108	0.144	0.159	0.045
12:00	0.014	0.093	0.114	0.310	0.108	0.144	0.159	0.040
12:15	0.014	0.093	0.114	0.310	0.107	0.144	0.159	0.040
12:30	0.046	0.099	0.135	0.257	0.107	0.118	0.151	0.053
12:45	0.046	0.099	0.135	0.257	0.107	0.118	0.151	0.053
13:00	0.046	0.099	0.135	0.257	0.107	0.118	0.151	0.053
13:15	0.046	0.099	0.135	0.257	0.138	0.118	0.151	0.053
13:30	0.047	0.125	0.163	0.242	0.138	0.106	0.182	0.056
13:45	0.047	0.125	0.163	0.242	0.138	0.106	0.182	0.056
14:00	0.047	0.125	0.163	0.242	0.138	0.106	0.182	0.054
14:15	0.047	0.125	0.163	0.242	0.155	0.106	0.182	0.054
14:30	0.023	0.114	0.169	0.258	0.155	0.102	0.231	0.050
14:45	0.023	0.114	0.169	0.258	0.155	0.102	0.231	0.050
15:00	0.023	0.114	0.169	0.258	0.155	0.102	0.231	0.044
15:15	0.023	0.114	0.169	0.258	0.113	0.102	0.231	0.044
15:30	0.043	0.142	0.183	0.314	0.113	0.118	0.210	0.081
15:45	0.043	0.142	0.183	0.314	0.113	0.118	0.210	0.081
16:00	0.043	0.142	0.183	0.314	0.113	0.118	0.210	0.081
16:15	0.043	0.142	0.183	0.314	0.105	0.118	0.210	0.066
16:30	0.032	0.161	0.198	0.315	0.105	0.142	0.168	0.084
16:45	0.032	0.161	0.198	0.315	0.105	0.142	0.168	0.084
17:00	0.032	0.161	0.198	0.315	0.105	0.142	0.168	0.084
17:15	0.032	0.161	0.198	0.315	0.118	0.142	0.168	0.084
17:30	0.040	0.169	0.216	0.268	0.118	0.180	0.162	0.087
17:45	0.040	0.169	0.216	0.268	0.118	0.180	0.162	0.087
18:00	0.040	0.169	0.216	0.268	0.118	0.180	0.162	0.080
18:15	0.040	0.169	0.216	0.268	0.168	0.180	0.162	0.080
18:30	0.081	0.191	0.258	0.308	0.168	0.205	0.240	0.161
18:45	0.081	0.191	0.258	0.308	0.168	0.205	0.240	0.161
19:00	0.081	0.191	0.258	0.308	0.168	0.205	0.240	0.161
19:15	0.081	0.191	0.258	0.308	0.215	0.205	0.240	0.161
19:30	0.134	0.164	0.289	0.320	0.215	0.165	0.335	0.186
19:45	0.134	0.164	0.289	0.320	0.215	0.165	0.335	0.186
20:00	0.134	0.164	0.289	0.320	0.215	0.165	0.335	0.186
20:15	0.134	0.164	0.289	0.320	0.225	0.165	0.335	0.186
20:30	0.104	0.232	0.286	0.310	0.225	0.137	0.379	0.193
20:45	0.104	0.232	0.286	0.310	0.225	0.137	0.379	0.193
21:00	0.104	0.232	0.286	0.310	0.225	0.137	0.379	0.182
21:15	0.104	0.232	0.286	0.310	0.243	0.137	0.379	0.182
21:30	0.046	0.214	0.255	0.336	0.243	0.133	0.331	0.133
21:45	0.046	0.214	0.255	0.336	0.243	0.133	0.331	0.143
22:00	0.046	0.214	0.255	0.336	0.243	0.133	0.331	0.140
22:15	0.046	0.214	0.255	0.336	0.225	0.133	0.331	0.207
22:30	0.035	0.174	0.227	0.307	0.225	0.103	0.249	0.208
22:45	0.035	0.174	0.227	0.307	0.225	0.103	0.249	0.156
23:00	0.035	0.169	0.227	0.307	0.225	0.097	0.249	0.156
23:15	0.035	0.169	0.227	0.307	0.203	0.097	0.249	0.180
23:30	0.041	0.104	0.208	0.313	0.203	0.086	0.216	0.180
23:45	0.041	0.104	0.211	0.303	0.203	0.086	0.213	0.180
Total kWh/Day	1.0757	3.160891	4.519056	7.136191	3.775	2.862913	4.977483	2.770354

APPENDIX D: Feeder section used for validation of feeder loss estimation



6th Lane, Mission Road,

Pita-Kotte, Sri Lanka

APPENDIX E : Matlab Program for Feeder Loss Validation

```
Rx1=0.028;
Rx2=0.028;
Rx3=0.028;
Rx4=0.038;
Ry1=0.032;
Ry2=0.032;
Ry3=0.032;
Ry4=0.044;

for p=1:96
 Vx(p,:)=[230,230*complex(-sin(pi/6),cos(pi/6)),230*complex(-
sin(pi/6),-cos(pi/6)),0];
 Vy(p,:)=[230,230*complex(-sin(pi/6),cos(pi/6)),230*complex(-
sin(pi/6),-cos(pi/6)),0];

    for k=1:10

    Ix(p,1)=conj(complex(0.8*Sx(p,1),0.6*Sx(p,1))/Vx(p,1));
    Ix(p,2)=conj(complex(0.8*Sx(p,2),0.6*Sx(p,2))/Vx(p,2));
    Ix(p,3)=conj(complex(0.8*Sx(p,3),0.6*Sx(p,3))/Vx(p,3));
    Ix(p,4)=-(Ix(p,1)+Ix(p,2)+Ix(p,3));
    Iy(p,1)=conj(complex(0.8*Sy(p,1),0.6*Sy(p,1))/Vy(p,1));
    Iy(p,2)=conj(complex(0.8*Sy(p,2),0.6*Sy(p,2))/Vy(p,2));
    Iy(p,3)=conj(complex(0.8*Sy(p,3),0.6*Sy(p,3))/Vy(p,3));
    Iy(p,4)=-(Iy(p,1)+Iy(p,2)+Iy(p,3));

    Vx(p,1)=V(p,1)-Rx1*(Ix(p,1)+Iy(p,1));
    Vx(p,2)=V(p,2)*complex(-sin(pi/6),cos(pi/6))-
Rx2*(Ix(p,2)+Iy(p,2));
    Vx(p,3)=V(p,3)*complex(-sin(pi/6),-cos(pi/6))-
Rx3*(Ix(p,3)+Iy(p,3));
    Vx(p,4)=V(p,4)-Rx4*(Ix(p,4)+Iy(p,4));
    Vy(p,1)=Vx(p,1)-Ry1*Iy(p,1);
    Vy(p,2)=Vx(p,2)-Ry2*Iy(p,2);
    Vy(p,3)=Vx(p,3)-Ry3*Iy(p,3);
    Vy(p,4)=Vx(p,4)-Ry4*Iy(p,4);

    end

Loss(p,1)=abs(Ix(p,1)+Iy(p,1))^2*Rx1+abs(Ix(p,2)+Iy(p,2))^2*Rx2+abs(
Ix(p,3)+Iy(p,3))^2*Rx3+abs(Ix(p,4)+Iy(p,4))^2*Rx4+abs(Iy(p,1))^2*Ry1
+abs(Iy(p,2))^2*Ry2+abs(Iy(p,3))^2*Ry3+abs(Iy(p,4))^2*Ry4;

end
```

APPENDIX F: Voltage profile for feeder starting point

	Phase 1 (V)	Phase 2 (V)	Phase 3 (V)
0:00	230	233	232
0:15	229	232	231
0:30	229	233	232
0:45	229	234	233
1:00	230	233	232
1:15	230	233	233
1:30	230	234	233
1:45	231	234	232
2:00	231	235	234
2:15	229	234	233
2:30	228	233	232
2:45	229	233	233
3:00	229	233	233
3:15	229	233	233
3:30	229	234	233
3:45	228	233	232
4:00	227	233	232
4:15	227	232	231
4:30	226	231	230
4:45	227	232	231
5:00	226	232	230
5:15	224	231	230
5:30	226	231	231
5:45	227	232	229
6:00	226	232	229
6:15	226	231	229
6:30	227	230	229
6:45	228	231	230
7:00	228	231	229
7:15	227	232	230
7:30	227	231	229
7:45	229	232	230
8:00	230	232	231
8:15	227	230	229
8:30	228	231	230
8:45	227	231	230
9:00	226	230	229
9:15	227	230	229
9:30	228	232	232
9:45	229	231	231
10:00	228	230	231
10:15	229	231	231
10:30	230	232	231
10:45	228	232	231
11:00	228	231	232
11:15	228	231	231
11:30	227	231	230

11:45	227	231	231
12:00	228	231	231
12:15	229	232	232
12:30	231	233	232
12:45	230	233	232
13:00	229	233	232
13:15	230	233	232
13:30	231	233	232
13:45	230	233	231
14:00	229	233	230
14:15	228	230	229
14:30	227	230	229
14:45	226	229	228
15:00	226	230	228
15:15	227	232	230
15:30	228	233	231
15:45	228	233	231
16:00	228	233	231
16:15	228	233	232
16:30	229	233	232
16:45	231	234	234
17:00	230	234	233
17:15	231	235	234
17:30	229	233	230
17:45	228	232	230
18:00	229	233	230
18:15	230	234	233
18:30	228	232	231
18:45	227	231	231
19:00	227	232	230
19:15	228	232	230
19:30	228	232	230
19:45	228	232	230
20:00	228	232	230
20:15	229	233	230
20:30	228	233	230
20:45	228	233	231
21:00	229	234	231
21:15	228	233	231
21:30	227	232	232
21:45	226	232	230
22:00	227	231	230
22:15	226	232	231
22:30	229	232	231
22:45	227	231	231
23:00	228	232	232
23:15	228	233	233
23:30	228	233	232
23:45	229	233	233

APPENDIX G : Actual Load profiles for 81 customers (in Watt)

Customer No:	0:00	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00	3:15	3:30
1	125	112	112	109	109	107	108	109	104	99	102	104	106	114	130
2	72	62	62	62	62	70	70	40	40	50	50	70	70	50	50
3	287	270	270	252	252	240	238	233	218	218	222	233	240	220	213
4	222	208	208	217	217	202	205	223	162	162	160	145	140	138	134
5	150	150	110	180	180	180	170	170	120	100	100	110	130	140	120
6	240	228	214	192	186	188	192	188	176	166	159	172	189	180	173
7	626	891	891	851	851	831	831	831	541	541	524	524	673	540	580
8	73	69	69	72	72	72	72	69	73	80	78	80	80	75	75
9	259	271	271	258	258	257	257	264	261	261	261	263	247	253	250
10	159	142	143	134	134	137	143	163	147	143	141	144	131	117	127
11	78	77	77	67	84	91	73	68	59	73	45	41	75	60	48
12	145	127	136	149	149	137	164	180	160	160	160	170	160	137	149
13	189	286	286	230	335	240	300	360	278	290	225	240	333	240	293
14	190	238	220	211	253	224	211	183	174	222	185	185	185	274	192
15	178	187	187	149	149	158	156	153	89	89	96	98	88	101	96
16	97	84	84	53	53	36	36	36	30	30	30	30	40	33	33
17	82	87	112	100	100	60	60	105	88	48	72	77	120	125	110
18	158	160	70	133	133	123	140	90	40	85	85	85	205	80	80
19	263	253	253	260	260	269	256	269	280	300	297	257	278	293	253
20	87	84	84	100	100	80	67	67	80	80	80	80	80	55	55
21	131	143	143	130	120	134	129	117	129	125	130	190	174	178	162
22	830	1060	823	843	1120	760	760	800	800	870	1140	1160	1020	960	773
23	71	63	63	65	65	58	65	63	48	60	74	57	57	47	57
24	198	198	210	210	210	189	188	188	197	197	207	207	358	358	377
25	99	99	106	106	106	103	110	100	93	93	140	120	113	99	99
26	151	151	148	147	147	132	127	142	142	142	147	148	141	136	136
27	115	120	120	90	90	110	133	100	60	100	93	73	87	140	160
28	232	212	220	204	228	260	220	172	184	212	224	228	220	188	200
29	159	154	147	144	149	149	154	154	132	132	125	125	135	135	110

Rest of the data are burnt in CD as Appendix G

APPENDIX H : Normalized load profiles for 81 customers (in Watt)

Customer No:	0:00	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00
1	27	24	24	24	24	23	23	24	23
2	32	28	28	28	28	32	32	18	18
3	48	45	45	42	42	40	40	39	36
4	99	93	93	97	97	90	91	100	72
5	41	41	30	49	49	49	46	46	33
6	58	55	52	46	45	45	46	45	42
7	46	66	66	63	63	62	62	62	40
8	30	28	28	29	29	29	29	28	30
9	33	34	34	32	32	32	32	33	33
10	61	55	55	52	52	53	55	63	57
11	23	23	23	20	25	27	21	20	17
12	38	33	36	39	39	36	43	47	42
13	35	53	53	43	62	44	56	67	52
14	41	52	48	46	55	49	46	40	38
15	51	53	53	42	42	45	44	44	25
16	41	36	36	23	23	15	15	15	13
17	24	25	32	29	29	17	17	30	25
18	66	67	29	56	56	52	59	38	17
19	46	44	44	45	45	47	45	47	49
20	119	116	116	137	137	110	92	92	110
21	28	31	31	28	26	29	28	25	28
22	55	71	55	56	75	51	51	53	53
23	23	20	20	21	21	18	21	20	15
24	49	49	52	52	52	47	47	47	49
25	26	26	27	27	27	27	28	26	24
26	35	35	35	34	34	31	30	33	33
27	31	32	32	24	24	30	36	27	16
28	50	46	48	44	50	56	48	37	40
29	40	39	37	37	38	38	39	39	33
30	11	9	12	17	17	3	3	3	16
31	39	26	28	28	28	29	26	25	31
32	46	41	37	37	45	45	38	38	33
33	33	33	31	20	20	15	10	9	2
34	0	0	0	0	0	0	0	0	0
35	45	44	43	41	43	39	39	41	38
36	29	33	33	29	29	30	29	29	29

Rest of the data are burnt in CD as Appendix H

APPENDIX I

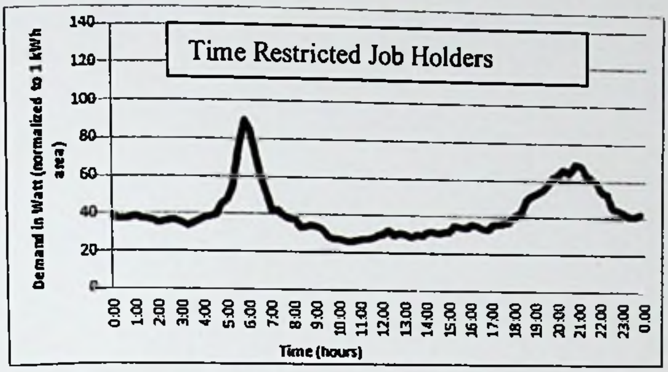


Figure 4.9 Cluster 1 Representative Curve

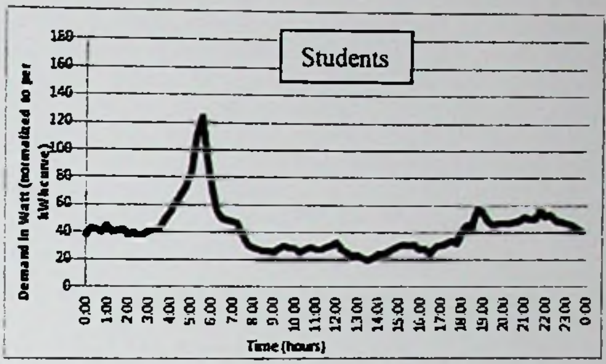


Figure 4.17 Cluster 9 Representative Curve

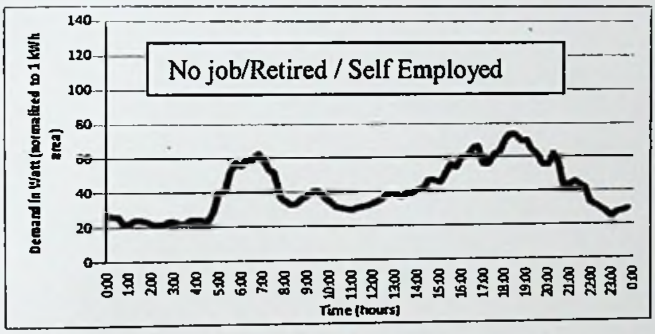


Figure 4.15 Cluster 7 Representative Curve

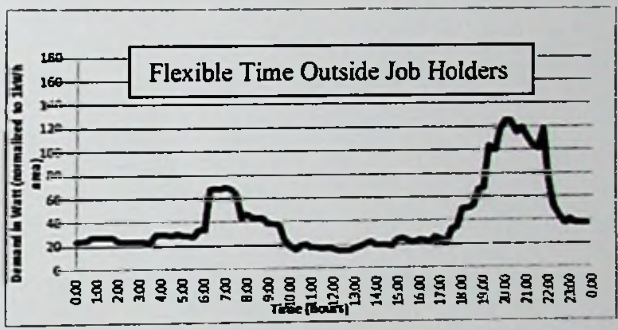


Figure 4.21 Cluster 12 Representative Curve

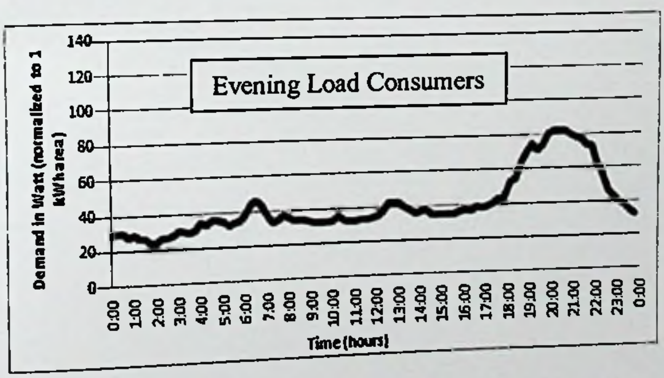


Figure 4.11 Cluster 2 Representative Curve

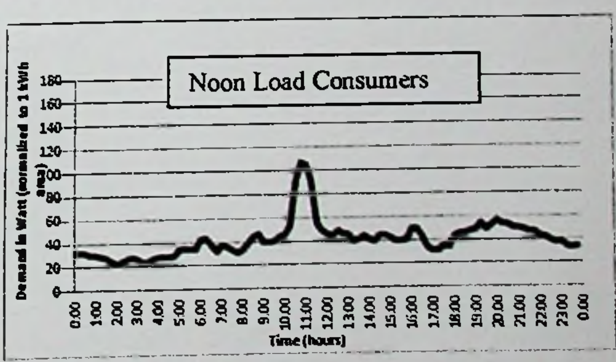


Figure 4.13 Cluster 6 Representative Curves

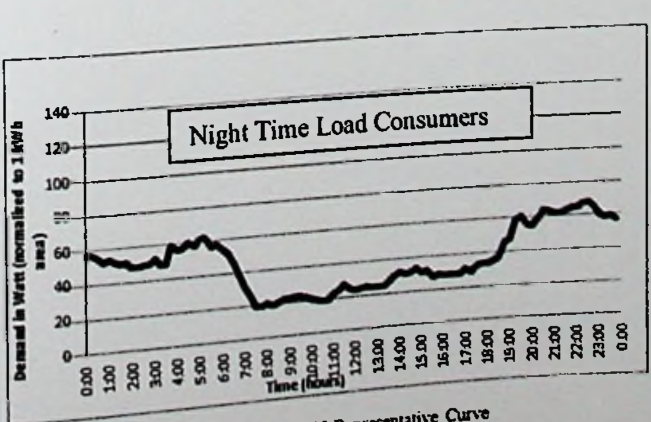


Figure 4.19 Cluster 10 Representative Curve

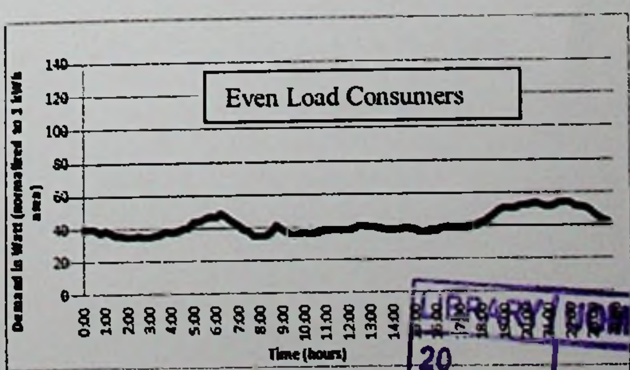


Figure 4.23 Cluster 13 Representative Curve