FORECASTING CUSTOMER ELECTRICITY LOAD DEMAND IN THE POWER TRADING AGENT COMPETITION USING MACHINE LEARNING

SAIFUL ABU

Master's Program in Computer Science

APPROVED:

Christopher Kiekintveld, Chair, Ph.D.

M. Shahriar Hossain, Ph.D.

Paras Mandal, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

to my

MOTHER and FATHER

with love

FORECASTING CUSTOMER ELECTRICITY LOAD DEMAND IN THE POWER TRADING AGENT COMPETITION USING MACHINE LEARNING

by

SAIFUL ABU

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Master's Program in Computer Science

THE UNIVERSITY OF TEXAS AT EL PASO

August 2016

Acknowledgements

At first, I would like to thank advisor, Dr. Christopher Kiekintveld of the Computer Science Department for everything he has done for me. I was very new to the research area and found the early days quite depressing. All my attempts to solve the problem failed for the first six months. I gave up. But he did not give up on me. He set aside time for the weekly meeting to know my progress. He listened to my works and results. I found him a reliable person to share my research and personal issues. If he was not there, I could not have finished my thesis.

I also want to thank my other committee members, Dr. Shahriar Hossain of the Computer Science Department and Dr. Paras Mandal of the Electrical Engineering Department. Their guidance and constructive criticism helped me improve my thesis.

Additionally, I want to thank the following persons individually for what they have done for me:

Dr. Olac Fuentes

> I worked with Dr. Fuentes as teaching assistant (TA) for the course algorithms and data structures. He made the TAs to attend the classes. I attended all his classes. I am strongly influenced by his teaching. He is very detail oriented. He is strict at the same time supportive. The classes helped me learn the concepts related to algorithms and data structures. With his support, an extreme introvert like me was able to run the lab section for the course and was able to develop public relational skills. I express my deepest gratitude to him.

Oscar Veliz

> I am forever in debt to Oscar, my colleague in the IASRL lab. Though he is a busy PhD student, he found time to go through my initial and sloppy thesis

draft. I found him very hard working, helpful and encouraging. He always had good suggestions for all my queries. I regret that I did not have more conversation during my stay in the lab.

Dr. Ivan Gris

Ivan is one of the smartest persons I have ever seen. He is incredibly practical and supportive. I received numerous suggestions and advice from him. I found him very reliable and hard working. His mere presence positively influences a crowd around him.

My Roommates

I received a tremendous amount of support from my two roommates Porag and Sunny. They were peaceful. Porag reviewed my draft thesis several times. I thank them for all the things they did to me.

And finally I want to thank my wife for giving me all the supports along the way.

NOTE: This thesis was submitted to my Supervising Committee on the July 21, 2016.

Abstract

Accurate electricity load demand forecasting is an important problem in managing the power grid for both economic and environmental reasons. The Power TAC simulation provides a platform for research on smart grid energy generation and distribution systems. Brokers are the focus of the design task posed to developers by the system. The brokers work as self-interested entities that try to maximize profits by trading electricity across multiple markets. To be successful, a broker has to forecast the electricity demand for customers as accurately as possible so it can use this information to operate efficiently. My proposed forecasting method uses a combination of clustering and classifiers. The customers are clustered based on a small history of weekly average load. After that, energy load history and weather related information are used as features to train classifiers for each cluster of customers. To forecast for a new customer, the proposed method needs at least one week of energy load history for the customer. The system assigns the new customer to one of the clusters based on the similarity of its electricity usage history. The classifier for that cluster will be used to forecast the new customer. This approach produced 13% error compared to 31% relative absolute error observed for a moving average baseline predictor. Power TAC has six different types of customer such as customers with demand shifting capabilities, customers with no demand shifting capabilities, electric vehicles, thermal storage, wind production and solar production. Previous approaches to demand forecasting treated all types of customers equally. This work shows that a forecasting system that treats customers of different type differently by creating clusters of similar types can generalize effectively, having similar error rates to learning individual predictors for each cluster, while also allowing fast predictions for novel customers.

Table of Contents

List of Figures

Chapter 1

Customer Load Forecasting in Power Trading Agent Competition

We cannot think of modern civilization without the existence of electricity. Across the globe, electricity generation companies are spending trillions of dollars every year to produce Terawatt-hours of electricity to keep the wheel of the civilization running [1]. A major portion of the electricity is produced by burning coal and natural gas [1]. Alone the electricity generation is responsible for more than 40% of all the greenhouse gas emissions [33]. Around 4 million customers were disconnected from electricity in blackouts during 2003 [4]. Electricity generation companies need to accurately forecast electricity load demand to avoid negative impacts on the environment, economic loss, and blackouts [19, 31]. The companies may improve their load forecasting methods by utilizing the wealth of electricity usage data of their customers. In my thesis, I propose a data driven electricity load forecasting system. The system was tested on a realistic, smart grid [7] simulation platform called the Power Trading Agent Competition (Power TAC) [16]. The proposed mechanism is capable of making demand forecasts about a novel customer using data of known customer.

1.1 Smart Grid

In a power grid, there are three subsystems: electricity generation, transmission, and distribution [7]. In a typical electricity generation subsystem, the generator rotates a turbine in a magnetic field which generates electricity. Usually, the turbine rotates using the power of kinetic energy of water falling over a waterfall, a river with strong current, or from the reaction in nuclear power plant, or energy produced from burning coal or natural gas. The energy generation system then transmits the electricity through a transmission grid and finally through the distribution grid. A traditional generation system is one way, meaning that electricity the flow occurs only from generation node to consumption locations, and there is a little capability to collect information and intelligently manage the grid at all levels. In Figure 1.1, flow of electricity is shown by an arrow. It can be seen that all the arrows start from electricity generation and in the consumer's direction which indicates flow of electricity.

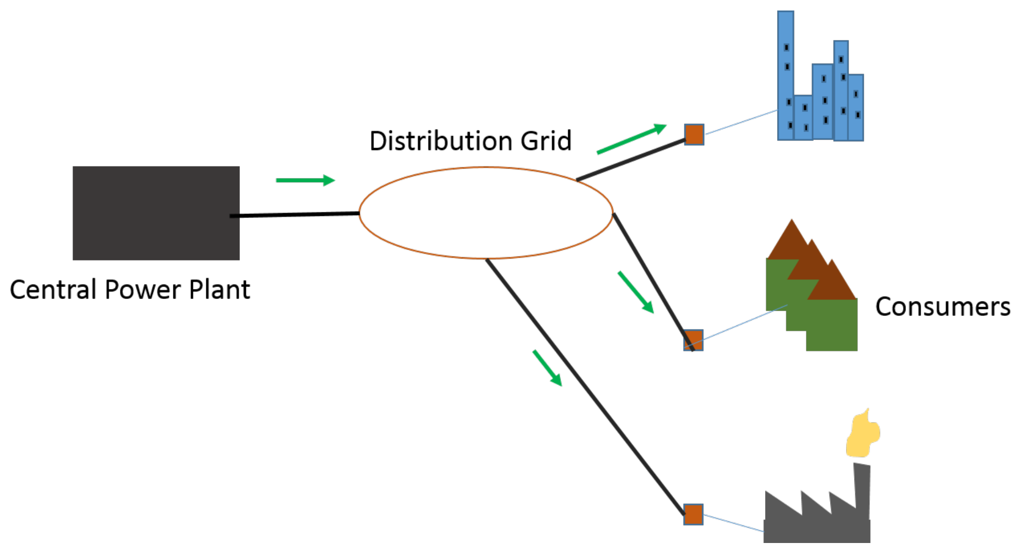

Figure 1.1: Traditional electricity generation and distribution system.

In contrast to the traditional electricity generation system, the Smart Grid (SG) is two-way [7]. Any node in the distribution grid can both consume and produce electricity to supply to the distribution grid. In Figure 1.2 a smart grid based electricity generation and distribution system are depicted. We can see that, in this system a flow toward the distribution system is possible. For example, renewable energy sources and household customers with solar energy production capabilities are pushing electricity to the power grid. The National Institute of Standards and Technology report states that the SG could make the electricity generation and supply robust against generator or distribution node failure, promote the use renewable energy, widely and efficiently, reduce greenhouse gas emissions, reduce oil consumption by encouraging usage of electric vehicles, and give customers more freedom to choose among energy sources. Smart grids will encourage usage of electric vehicles since these vehicles have the ability to store power in a battery and transmit the power to the distribution grid if there is a need [7]. Customers are showing strong motivation to use renewable energy as indicated by the statistics that 20% of total energy is from the renewable sources which are second after coal 24%. Consumers are using renewable energy due to economic reward and environmental concern. The major challenge with the usage of renewable energy as part of the supply is that it is uncertain [29]. This uncertainty causes problems for predicting how much energy the SG can produce in a future time slot. To manage the SG efficiently we will need effective methods to predict both energy production and load demand so they can be balanced in real time [26].

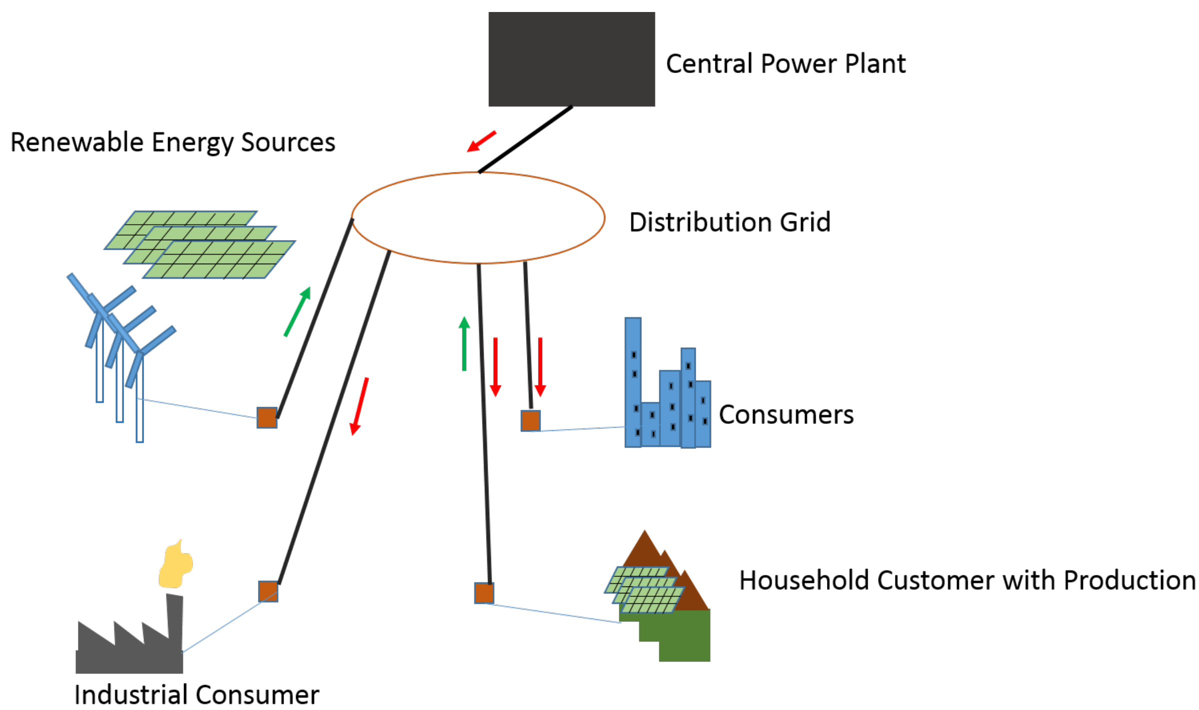

Figure 1.2: Smart grid based electricity generation and distribution system.

1.2 The Power Trading Agent Competition

The Power Trading Agent Competition (Power TAC), is an international research competition based on a smart grid energy market and distribution system [16, 17]. The Power TAC simulation models components including a wholesale market, electricity brokers, customers, a distribution utility and a weather service [17]. The brokers publish tariff plans for electricity consumers and producers. They buy electricity from the wholesale and balancing market to meet net customers demand. The system has realistic customer models and uses real weather data. Because of the realistic customer models present in the system, Power TAC can be used to do research on the customer electricity load demand. It is also useful because generating data is cheap in a computer simulation. The following sections give a brief explanation of each component of the Power TAC. Figure 1.3 shows a block diagram of the components of the Power TAC simulation environment.

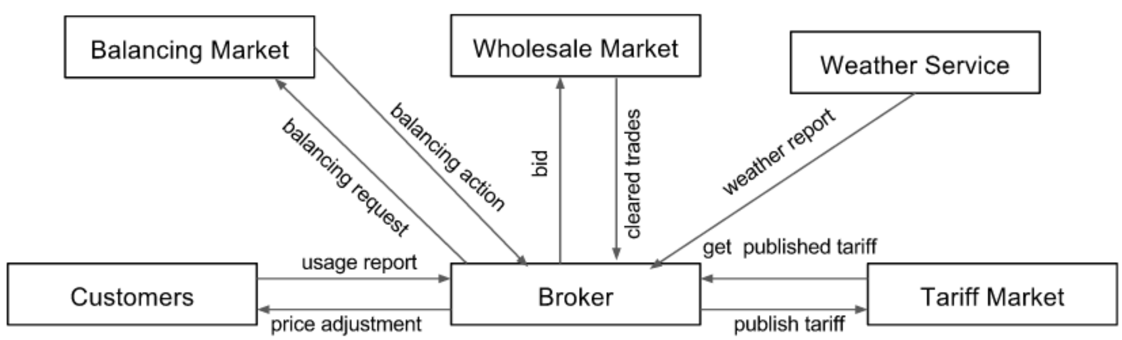

Figure 1.3: PowerTAC simulation environment.

1.2.1 Markets and Distribution Utility

There are three different types of markets in the Power TAC simulation, namely, the wholesale market, tariff market and balancing market. The wholesale market is the bidding place for buying or selling energy. Bulk energy producers and brokers take part in the wholesale market auction. Brokers can submit their bids for 24 future timeslots in the wholesale market by specifying the price they are prepared to pay. If the bid was successful, the broker receives the desired quantity by paying the money. At each time slot, the system notifies the broker about the wholesale market clearing prices. Brokers publish their tariff plans in the tariff market. A tariff holds information about the pricing of the energy. Customers, upon analyzing available tariffs, subscribe to the most suited tariff plan. The balancing market represents the market from where the broker can buy energy in case of imbalance. For example, if a broker has bought less amount of energy for a given timeslot and it finds it needs more energy then it can buy the necessary amount of energy from the balancing market. Usually, the balancing market transactions are more costly for brokers than the wholesale market.

The distribution utility has two main objectives. First, it supplies energy to the consumers from the wholesale market and from the renewable producers. Secondly, it works as a default broker that publishes default tariffs at the start of the game. This makes sure no customer is ever out of energy. Brokers can publish more lucrative tariffs to attract customers.

1.2.2 Brokers

In Power TAC, participants implement a fully autonomous broker agent. A broker in the simulation competes with the other brokers by publishing competitive tariffs. The goal of a broker is to maximize its profit by buying electricity at a cheap price and selling the electricity at a profitable price. Here is the list of actions that a broker can take in Power TAC simulation:

- At any hour of the simulation, a broker can publish a new tariff. Each tariff is targeted to a specific category of customers. The tariffs contain information about which category of customers it is targeted, the expiration date of the tariff, signing bonus, the penalty for early withdraw, and the rate of periodic payment.

- At any time, a broker can modify its published tariffs. It can adjust payment rates and withdraw penalty. The broker can also revoke a tariff that is not profitable.

- To meet customer demand, a broker takes part in the electricity auction in the wholesale market. The wholesale market is the meeting place for bulk electricity sellers and brokers. The broker specifies how much electricity it needs, how much it is ready to pay for and at which time it needs the electricity. The broker can make asks or bid for next 24 hours. Based on the asks and bids from other brokers in the simulation, the broker's bid may or may not get cleared.

At any moment in the simulation the brokers have the following information:

- Participating brokers in the simulation

- Information such as population, power types of the customers present in the simulation.

- Initial electricity usage data of customers and initial clearing prices of the wholesale market.

- Published tariffs in the simulation.

- Information about tariff modifications or revocations.

- Wholesale market clearing prices for the last time slot.

- Energy transaction information of subscribed customers.

- Transaction in the balancing market.

- Current bank balance for the agent.

1.2.3 Customers

A customer represents an entity that buys energy from or sells energy to the brokers. A customer tries to minimize its costs by subscribing to a broker that offers profitable tariff for the customer. The customers occasionally evaluate the tariffs present in the market and switch to a profitable tariff if necessary. Customers decide the outcome of a competition as the customers are the only consumers present in the system. A customer has the following attributes:

- A unique name.
- Number of individuals the customer represents. This number can range from one to several thousand.
- A power type that specifies which category of customers (producer or consumer) the model falls into. Producer and consumer categories have several subcategories. In Power TAC the categories of the customers are consumption, interruptible consumption, thermal storage, wind production, solar production and electric vehicle. Each customer type has its distinctive behavior.

A customer can take the following actions during a simulation:

- Evaluate available tariffs in tariff market.
- Subscribe to or abandon a tariff. Customers try to maximize their economic gain so if there is a lucrative tariff in the market, a customer might try to subscribe to it.
- Generate meter reading based on produced or consumed energy. The system then sends this meter reading to the broker it subscribed to.
- Customers with demand shifting capabilities can shift their demand to a favorable time slot.

Customer models and load forecasting problem will be described in more detail in Chapter 3.

1.2.4 Weather Service

The weather service broadcasts weather forecast for future hours and a weather report for the current hour to the brokers. The weather report contains information such as wind speed, cloud cover, temperature, the day of week and month of week. Power TAC uses real weather data from the past to make the simulation more realistic. Brokers can use this information to forecast demand for the weather sensitive consumers. The weather information also makes it possible to forecast renewable energy production.

1.3 Importance of Accurate Demand Forecasting in Power TAC

A broker has to make bids and asks in the wholesale market. The amount of electricity it asks for depends on the demand forecast of its subscribed customers. If the broker fails to make an accurate demand forecast, it will not be able to ask for the proper amount of electricity. So it will end up asking more or less energy than the required amount in the wholesale market. In this case, the broker will have to buy energy from the balancing market at a higher price, or have to sell surplus electricity at a lower rate. As a result, it will face monetary losses. This unwanted scenario can be avoided through demand forecasting as accurate as possible. My thesis investigates ways to find a better demand forecasting mechanism.

1.4 Main Research Questions

The brokers in Power TAC get only 5 seconds of time to communicate with the server. Within the 5 seconds, it has to decide how much electricity the customers are going to use, decide on an ask with a profitable bid and make other decisions. It will be advantageous if there is a way to have a single forecasting model that gives the aggregated demand prediction for all customers in a single shot. So my first research question is:

> How effectively can we make a single demand forecasting model that make demand forecasting about all the customers in a single shot?

As I solve the forecasting problem by training machine learning models, and the accuracy of the model will depend on which features I choose. If I can identify the most predictive features, the system will have less error. The second research question focuses on this problem:

> What are the most relevant features for the demand forecasting model?

Not all learning algorithms are suitable for a given scenario. I will be looking for learning algorithms that can be trained fast, that produces demand forecast as accurate as possible and makes predictions quickly. My third research question is:

> Which learning algorithms are the most suitable for the Power TAC scenario?

The system should be robust enough to be able to make demand predictions about a novel customer. My fourth research question is:

> Can we develop an effective way to make a forecasting model trained on known customers and that can forecast the demand of a novel customer?

I also need to see how the training set size impacts the accuracy of the forecasting model. If more training data the accuracy improves, I will try to gather as much data as possible until there are diminishing returns. My fifth research question is:

> What is the impact of training set size on the accuracy of the prediction model?

1.5 Contribution

In previous work on Power TAC, all the customers were treated the same. This work shows that different customers should be treated differently. I showed a demand prediction scheme for consumption type customer might not be suitable for interruptible consumption customers. This work proposes a method to make demand forecasting about a novel customer. During the training phase, known customers are grouped based on their weekly average electricity usage. Data of the same group are combined to make forecasting model for that group. A novel customer is assigned to one of the groups based on its weekly average electricity usage. The forecasting model of that group will be used to make forecasting about the new customer. This scheme produces better forecasting than a moving average baseline forecasting scheme. My work also found that there is little correlation with the accuracy of the system and the size of the training dataset.

1.6 Overview of the Thesis

In the second chapter of the thesis I present related work on electricity demand forecasting. In chapter 3, I present information about the customers present in Power TAC. The information include customer types, their behaviors and descriptive statistics about them. In chapter 4, I present the work I did to answer the research questions and show the results. In chapter 5, I present the significance of the results and future work.

Chapter 2

Related Work

In this chapter, I described different methods of energy load forecasting for long term and short term in the literature. It is hard to know the state of the art electricity consumption mechanism in Power TAC as most of the researchers did not publish their demand forecasting mechanisms [18, 24, 34, 23]. Therefore, I mostly describe the work done for real world electricity demand forecasting mechanisms.

2.1 Variables in Electricity Demand

Studies have found that electricity demand is highly correlated with temperature [12, 9, 6]. The study in [9] was done in the region of Australia. It was found that in a lower temperature the customers tend to use heaters and in a higher temperature they tend to use coolers. As a result, the increase or decrease of temperature from a certain point will cause the consumption of electricity to increase. In the study [6], two demand forecasting models were proposed. One was univariate Auto Regressive Integrated Moving Average (ARIMA) and another one was univariate ARIMA model along with a temperature depended transfer function. The model with temperature variables did better forecasting than the one without the temperature variable. On the other hand, the study in [5] showed that inclusion of temperature variables in the forecasting model actually introduced more error in demand prediction. The aim of the study was to make forecasts about electricity usage in January based on the past five years training data using a Support Vector Machine (SVM). The reason behind getting more error after including temperature variable may be because during January the temperature did not change much and the inclusion might have caused overfitting.

Cloud cover and wind speed are also correlated with electricity demand [12, 30]. As cloud cover increases, the demand for light increases too. The increased lighting demand causes increased electricity demand. The period where cloud cover was low, the electricity demand was also low [12]. High-speed wind across wet walls helps cool houses. High-speed wind thus may cause reduced electricity demand due to reduced demand of air cooling [30].

In the survey article [8], the authors reported that customer electricity demand gets affected by the weekday and the month in consideration. The electricity load demand can get low or high based on the day of the week. The weekends usually have different load demand pattern than the weekdays. Based on the hour of a given day, the load demand can be higher or lower too. The season also showed an impact on electricity demand.

2.2 Electricity Load Forecasting Using Time Series Analysis

To make electricity load forecast, researchers have used statistical methods such as statistical average, ARIMA and exponential smoothing. Agent TACTEX'13 [35], the winner of the PowerTAC competition in 2013, used the statistical average to make electricity demand forecasting for an hour of a day of a week. In a week a customer has $24 * 7 = 168$ hours or slots where it can consume electricity. TACTEX'13 kept track of average usage of 168 weekly slots for each customer. To predict a future time slot, the agent would look at which weekly slot the future time slot would fall in. Then the agent used that weekly slot's average usage as the forecast of the future time slot. Cho et al., have used an ARIMA model for load forecasting [6]. The ARIMA model uses both moving average and auto regression to forecast the demand. To make a forecast about a future time slot, the auto regression model uses some previously observed time slots values based on its degree. Moving average scheme would use the average of all the known time series data points to make a prediction about a future time slot. In the study [3], a short term load demand was proposed that uses several ARIMA models. For the combination of weekday and temperature level, 16 short term load forecasting models were used. This scheme made better forecasting than a single ARIMA model. Others used modified Halt Winter Exponential Smoothing for demand forecasting. The modified exponential method was capaple of dealing with weekly and daily seasonality pattern present in the data [15].

2.3 Load Forecasting using Classification

Support Vector Machines (SVM) have proved to be an effective tool for load demand forecasting [32, 5]. In [5], the authors used SVM to forecast electricity demand of January based on past 5 years electricity consumption data. Separate SVM models for separate seasons were proposed. SVM model trained with data from January was able to make better load forecasting. Artificial Neural Network (ANN) is another favorite load forecasting mechanism among the researchers [14], [27], [13]. Quan et al., [27], used ANN to model Prediction Interval (PI) to forecast renewable energy forecasting. Izgi et al., [14] used ANN to determine the time horizon suitable for solar energy production. Mandal et al., [20] worked on electricity load demand forecasting using ANN. Parra [25] et al., used various machine learning techniques to make 24 hours ahead load forecast for the Power TAC simulation. They found that hour of week, weather related features such as temperature and cloud cover were influential to the electricity load demand. The forecasting modules made a low error while forecasting for the customers that showed regularity in their energy consumption behavior. The application of linear regression [21, 10] and Kalman Filtering [2] also appeared for demand forecasting in the literature.

2.4 Load Forecasting using Clustering

Clustering can be used to group consumers with similar electricity demand patterns [10]. McLoughlin et al., [21] applied clustering based on 6 months electricity usage of household consumers in Ireland. Application of clustering generated common load patterns called load profiles which were used to forecast future load demand. Cho [6] et al., noticed that customers can be categorized to improve the accuracy of demand forecasting. They manually clustered the customers in four groups namely, commercial, office, residential and industrial customers. For Power TAC environment, Wang et al., [36] proposed a broker that clusters the bootstrap data of customers and generates a linear regression classifier for demand predicton for each cluster. Hernandez et al., [10] used kMeans and Self Organizing Map to cluster industrial park's consumption in Spain to understand micro environments present in a larger environment.

2.5 Load Forecasting using Expert Systems

Expert systems predict electricity demand prediction using variables that are likely to affect electricity demand [28, 11]. This system then mimics a human operator's steps to forecast load. This load forecasting mechanism appeared applicable for short term load forecasting [28, 11, 22].

From the review of the literature, the importance of weather related variables such as temperature, cloud cover and wind speed is evident. Also, the hour of the day and day of the week are highly correlated with the load demand. A combination of machine learning classifiers and clustering algorithms appears to be a better idea. For the methodology of [25] it will take a large number of predictors for the simulation system. Also, those predictors will not work if the name of the customer is changed or a new customer is introduced as each predictor is hard coded with a specific customer. It sounds reasonable to cluster the data first and then train machine learning classifier for each cluster. This approach will improve generality. Instead of training only on bootstrap data as proposed in [36], a wealth of data generated from the simulations can be used to train the cluster. Since the clustering is done offline, the proposed approach will not suffer from the problem of having a time limit that the broker has to face if the cluster is trained during the competition. After the clustering is done, for each cluster, different machine learning classifiers can be trained to figure out which one performs the best. So, the broker will no longer stick to linear regression as in [36].

Chapter 3

Customer Description

In this chapter, I describe in more detail the customer models implemented in the Power TAC simulation system, and present some descriptive statistics about their behaviors as an initial analysis to indicate important features for learning.

3.1 Customer Categories

In the Power TAC simulation, a customer can be electricity consumer or producer based on its power type. A customer evaluates the tariff plans targeted for its power type and can look for the tariff that minimizes its cost. There are several types of customers in the Power TAC simulation such as consumption, interruptible consumption, thermal storage, solar production, wind production and electric vehicle. Each power type has its own characteristics. For example, interruptible consumption customers can shift their electricity demand to some off-peak hour, the solar production customers can produce energy based on the weather condition. As opposed to previous methods of demand forecasting, I argue that each category of customers should have different load forecasting mechanism based on their power types. One load forecasting method can be suitable for a category of customers while it may be unsuitable for other categories because each category behaves differently. I describe the characteristics of the customers below.

- Consumption: Customers with power type consumption are the most common customers. They use energy when they need it. They cannot shift their demand to a future timeslot. Usually, they have a regular electricity usage pattern. Often, they show a similar pattern for weekdays, and they have similar kind of usage pattern for the weekends.

- Interruptible Consumption: Interruptible consumption customers are smart enough to shift their energy demand in a timeslot where they can buy electricity at a reduced price. Because of this shifting capability, they don't show as regular a usage pattern as the consumption customers do.

- Thermal Storage: Thermal storage customers can store electricity and can supply the stored electricity to the grid depending on its charge level. During a day, their electricity usage in a day depends much on the energy they used in the last timeslot.

- Solar Production: The solar energy production customer's energy production depends on the cloud cover. They are highly likely to produce energy during the day time.

- Wind Production: Wind production customers generate energy from the wind. Their production varies with wind speed.

- Electric Vehicle: An electric vehicle customer represents one electric vehicle. Their usage of energy is quite irregular and hard to predict.

Before diving into the problem of identifying good forecasting methods, I found it useful to take a look at how the customers of different power types behave. I wrote a log extractor program that extracts all the electricity consumption and production data by all the customer for all of the timeslots. At the end of the game, it makes a report on normalized usage of all the customers. Normalized values are useful because even if the amount of energy usage among the customers varies, the pattern of usage can be captured through it and normalized usages of different customers can be plotted on the same graph. Figures 3.1 to 3.3 show normalized electricity demand or supply for Mondays.

Zero in the x axis means hour 12:00 am. From Figures 3.1 to 3.3, it is clear that some customer's electricity usage is higly correlated with the hour of the day. The customers of type consumption and solar energy are example of these types. The other customer type demands did not have strong correlation with the hour of a day. In both consumption and solar energy customers the consumption or production curve grows fairly smoothly until it reaches a peak point. After the reaching the peak, the production or consumption drops smoothly. The customers with types interruptible consumption, wind production, thermal storage and electric vehicle had irregular demand/supply pattern.

In Figure 3.4 to 3.6, normalized electricity usages during the week are shown. Hour 0 to 23 represents all the hours of Monday from 12:00 am to 11:00 pm. It appears that the electricity demand for consumption customer, interruptible consumption, thermal storage had repetitive demand patterns every day. The solar energy production customer showed repetitive production pattern. Moreover some customers such as the consumption type and the thermal storage type showed different pattern during the weekend. During the weekend they usually had lower energy demand than the weekdays. In the case of electric vehicle and wind energy production customers the demand/supply patterns were not regular. From these observations, I hypothesize that the consumption customer and the solar energy type customers can be forecasted with the most accuracy. Due to the irregular patterns of other customers, it will be harder to make accurate demand forecasts about them.

3.2 Statistics

In this section, I present some descriptive analysis of the customers available in the system to get an initial understanding of the typical behavior of the models.

- **Customer Vs PowerType**

  The Figure 3.7 shows number of customers of each power type for a typical Power TAC simulation game. The electric vehicle power type has the largest number of

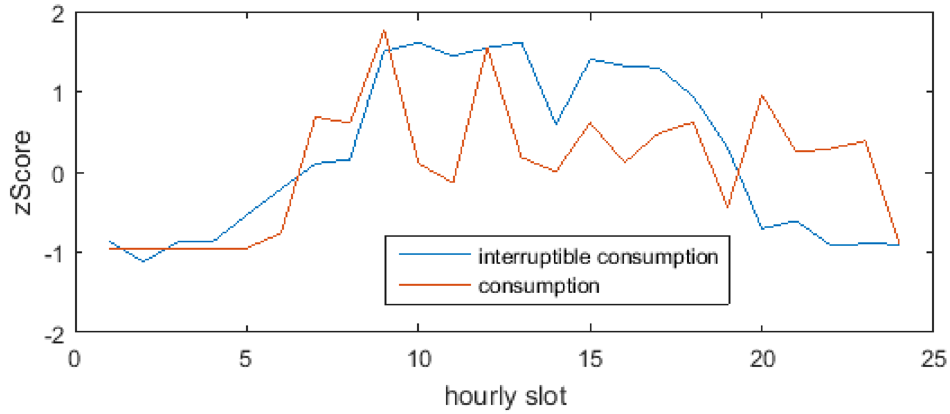

Figure 3.1: Electricity usage z Score over Monday

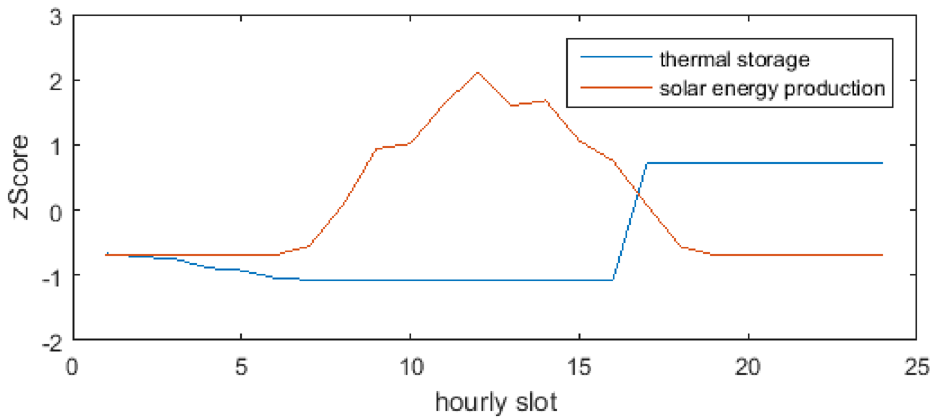

Figure 3.2: Electricity usage z Score over Monday

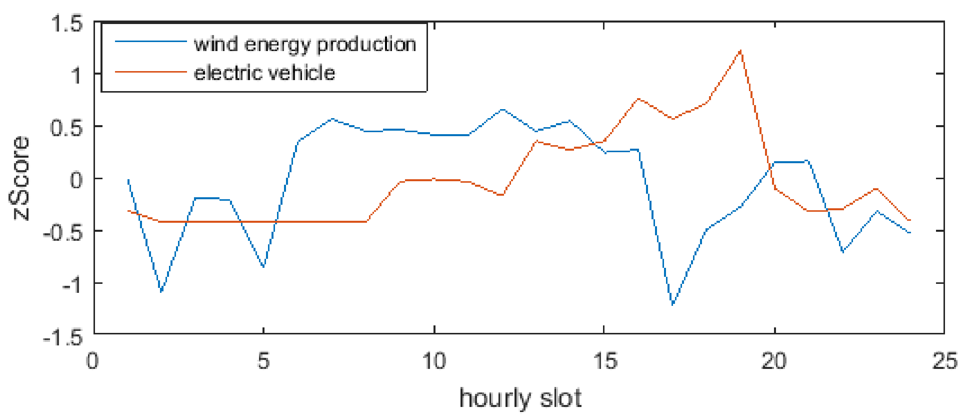

Figure 3.3: Electricity usage z Score over Monday

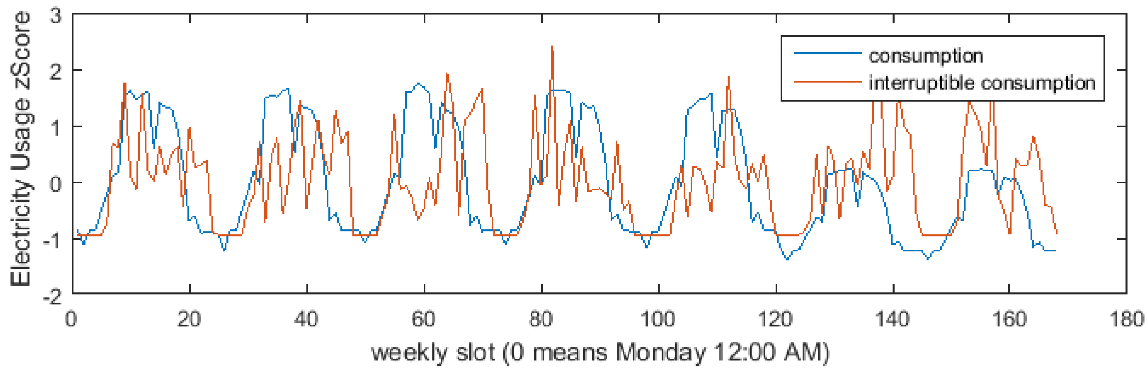

Figure 3.4: Electricity usage z Score over week

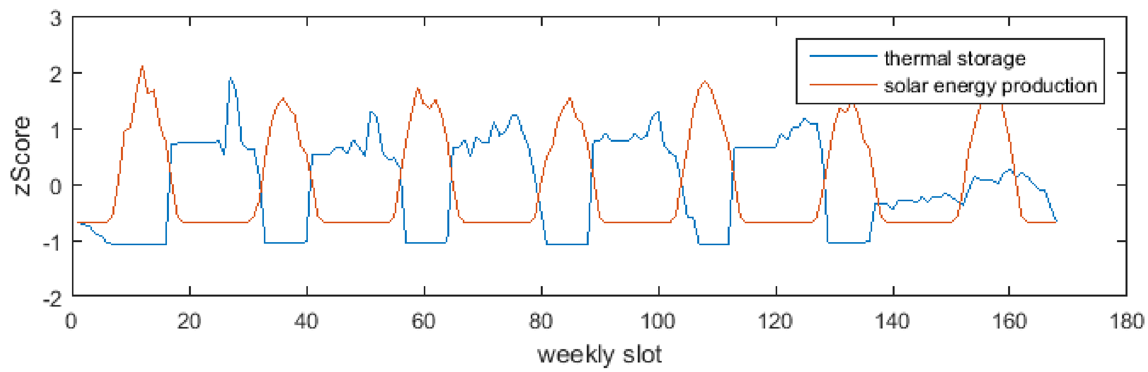

Figure 3.5: Electricity usage z Score over week

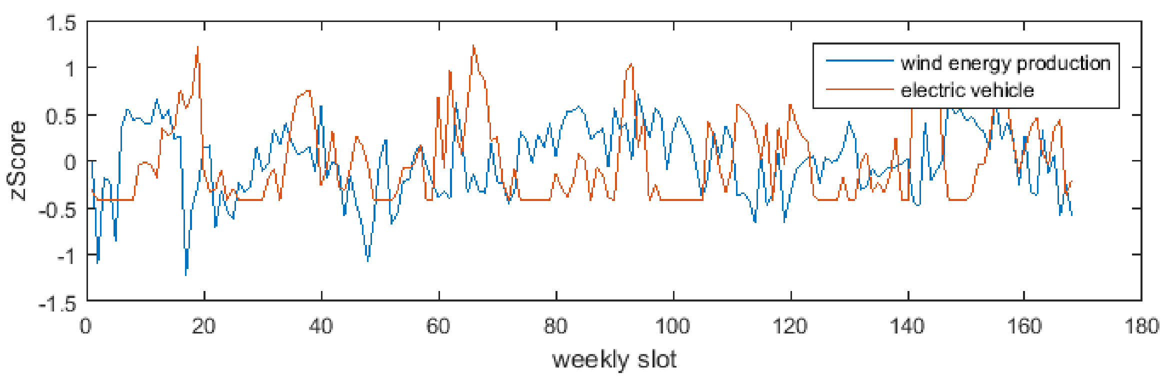

Figure 3.6: Electricity usage z Score over week

customers. This is because each electric vehicle represents a population of size 1. Consumption and interruptible consumption power type customers follow the electric vehicle power type customers in terms of number of customers.

- **Population Vs PowerType** From figure 3.8 the power type of consumption has by far the largest number of population. Some customers can represent thousands of individuals. For this reason, even though there are only a few numbers of consumption and solar energy customers, they can represent a population of several thousand individual customer households.

- **Total Energy Consumed Vs PowerType** The Figures 3.9 and 3.10 show the contribution to energy transactions by different power type customers in a typical Power TAC simulation. From the figure, we can see that the consumption type customers are responsible for the largest amount of energy transaction of total quantity (more than 60%). After the consumption type customers, the interruptible consumption type customers trade 18% of total traded energy. The solar production customers had 11% of the total energy transactions. A successful broker needs to forecast the consumption type customer with prime importance because of the bulk of energy they transact. Due to this fact, I concentrated mostly on forecasting about consumption type customer demand.

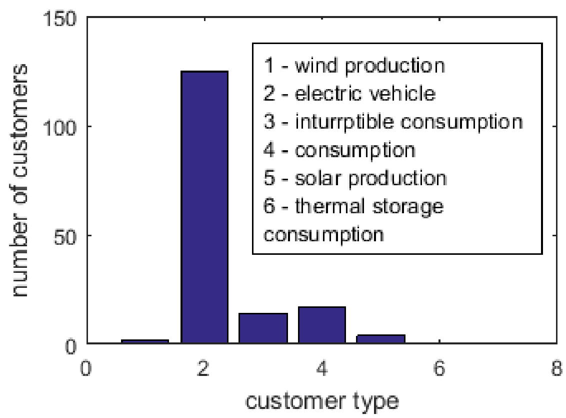

Figure 3.7:
Number of customers vs Powertype.

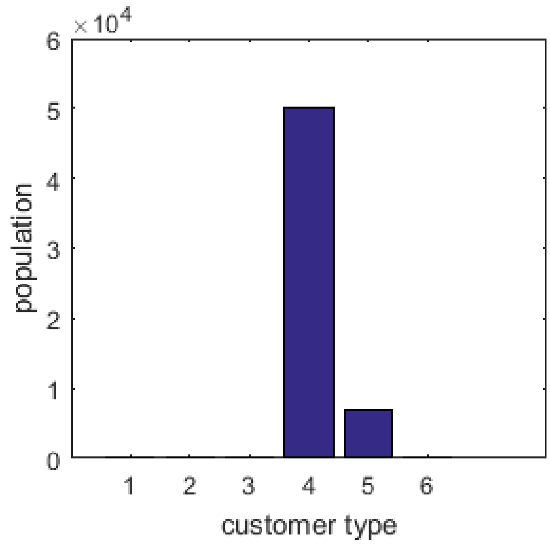

Figure 3.8: Population vs Powertype

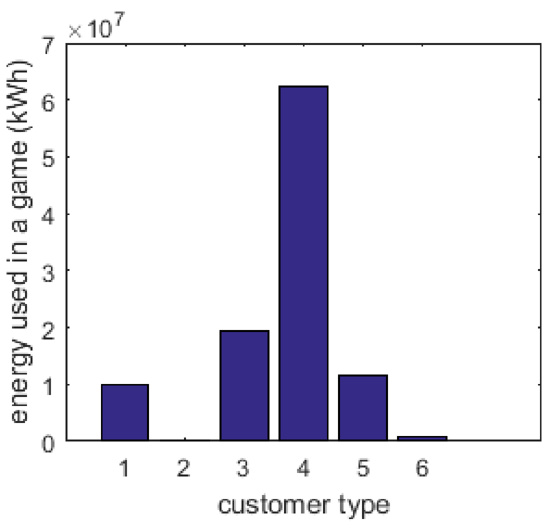

Figure 3.9: Energy vs PowerType.

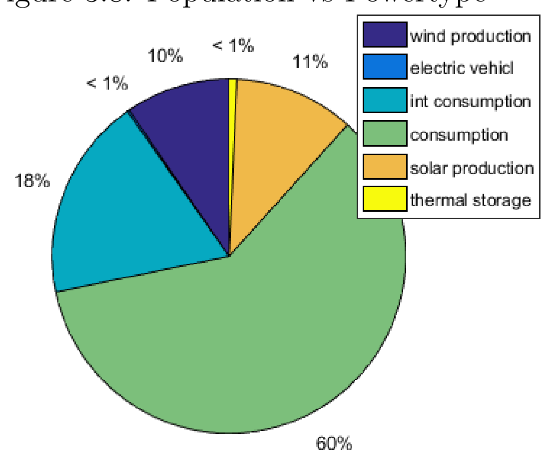

Figure 3.10:
Energy share for each power type.

Chapter 4

Methodology and Results

In this chapter, I describe the methodologies I followed to find answers to the research questions I have described in Chapter 1.

4.1 Learning Algorithms, Training, and Test Set

I have tried several learning algorithms that are available in Weka [37]. I was looking for learning algorithms that can be trained as quickly as possible with a large amount of data and with limited computing resources. The learning algorithm also has to produce a lower error rate and fast output during the time of competition. I have found that M5P, M5RULES, REP TREE, LINEAR REGRESSION [37] met my requirements. For training the models, I have used 10 game log files. Each log files contain about 10,000 training instances. A training instance contains the values of independent and dependent variables values. In the case of my thesis, example of independent variables would be temperature, cloud cover etc. and the dependent variable would be actual electricity usage of customers. Training time was getting more than 48 hours when I was using more than 10 log files. I separated 5 log files randomly from set of available game log files to use as a test set. The log files included in the test set were not used in training.

4.2 The Baseline Electricity Forecasting Mechanism

The first baseline energy forecasting mechanism is the default prediction mechanism provided by the Power TAC system. It exploits the fact that usage of a timeslot of a customer in a specific date is highly correlated with the day of a week and hour of a day. To make a prediction it stores the average energy usage of an hour of a week. So, for each customer, it uses $24 * 7 = 168$ values to remember average usages. As soon as it is informed of the new usage information of an hour of a week, it updates the old average using the Algorithm 1. To make demand forecasts about a customer it uses Algorithm 2.

Algorithm 1 Update average usage for $customer_i$ for day d and timeslot t, $newUsage$

1: avgUsage = get average usage of $customer_i$ at day d and time slot t

2: $avgUsage = 0.7 * avgUsage + 0.3 * newUsage$

Algorithm 2 forecast usage for day d and timeslot t for $customer_i$

1: avgUsage = get average usage of $customer_i$ at day d and time slot t

2: return $avgUsage$

4.3 Feature Extraction

All the activities that occurred in a game can be found in a game log. Activities such as buying or selling electricity occur during a time slot. At the beginning of a time slot, the system notifies the broker that a new time slot is about to begin. The system also notifies the brokers with weather forecast about the future time slots. As a time slot ends, the broker receives information about its customer's energy usage which is called tariff transaction report. From game log file, a feature extractor program can extract features related to customer time-related information, electricity usage, and weather variables. By keeping the previous electricity usage history for each customer, it is also possible to generate useful statistics such as mean and standard deviation usage for each hour of a weekday. Algorithm 3 specifies how the extraction program retrieves the necessary information from the tariff transaction report. As the broker gets notification of the beginning of a new time slot, the extraction program has all the information related to energy usage and weather data of the previous time slot.

Algorithm 3 extract information from transactionReport sent to broker after each time slot through TariffTransactionHandler call back method

```
timeSlot = get time slot from transactionReport
customerName = get customer name from transactionReport
energyUsed = get energy used from trom transactionReport
addUsage(customerName, timeSlot, energyUsed)
```

Algorithm 4 write extracted data after timeSlot update message received from TimeSlotUpdateHandler call back method

```
knownTimeSlot = timeSlot - 1
for each customer do
    day = get day of knownTimeSlot
    hour = get hour of knownTimeSlot
    statisticalData = get statistics of the customer of day and hour
    weatherData = get weather data of knownTimeSlot
    trueUsage = get true usage of customer in knownTimeSlot
    trainingInstance = create training instance by combining statistical data, weather
        data and true usage
    writeToFile(trainingInstance)
end for
```

4.4 Single Demand Predictor

The single demand predictor is designed to make a single aggregated load demand forecast for all the customer in a single shot. It is trained with 54 features. The features included time-related features such as day of week, hour of day, month of year; weather-related features such as cloud cover, wind speed, temperature; historical data such as past 24 slot's electricity usage, last week's electricity usage and statistical features such as average and standard deviation of the weekly slot. Algorithm 5 shows how a single prediction model is created. At first, features extracted from all types of customers are combined to make training set. Several classifiers are then trained based on the training set using 10 fold cross-validation [37]. It can be seen that the training set had information about all the customers of all types discussed in Chapter 3. For this scheme, the baseline predictor produced on average 70% root relative percentage error [37]. And for all learning algorithms, the error rate was more than 80%. The reason behind this poor performance is irregularity in training data. As we saw earlier, not all customers have regularity in their electricity usage pattern. As I have aggregated information about all the customers, which included irregularity, the training model is not able to generalize knowledge from the training set properly using the features and data available.

Algorithm 5 find a single best classifier

1: combine all slot based training instance of all the customers
2: train available classifiers on the combined data using 10 fold cross validation
3: choose the classifier with minimum error
4: save the classifier for making prediction about all the customers

4.5 Individual Demand Predictor

From the experience of a single demand predictor, I included only information about consumption customer data in training set and excluded information about interruptible consumption, thermal storage and electric vehicles power type customers. Also, this time I reduced the number of features using Weka's attribute selection algorithm [37]. Out of 54 features, only 5 were choosen. The features selected were temperature, wind speed, cloud cover, average of the weekly hour slot, standard deviation of the weekly hour slot. Features extracted for each customer served as a separate training data set. I trained an individual predictor model for each customer model. In general, if there are n customers in the system, we will need n electricity demand predictors, each trained on the data of a single customer. I went further by checking different machine learning algorithms such as M5Tree, Linear Regression, M5P rules and REP tree for each customer and picking the best performing one for each customer [37]. The process of finding best prediction model for each customer is shown in Algorithm 10. Each prediction model will make an electricity load demand prediction for each customer. The broker can calculate the total electricity demand by summing over all the predictions.

Algorithm 6 find best classifiers created for each individual customer

```
for each customer do
    combine all slot based training instance of the customer
    train available classifiers on the combined data using 10 fold cross validation
    choose the classifier with minimum error
    save the classifier for making prediction about the customer
end for
```

4.6 Cluster Based Demand Predictor

One problem with the individual demand predictor is that this system has a demand prediction model for each known customer. So, it cannot make demand forecasts for a novel customer. This is a serious problem because during a competition the system may introduce a new customer or change the name of a customer. In those cases the individual demand prediction mechanism becomes unreliable. To mitigate this issue, I propose a solution that groups similar customers based on their weekly average usage pattern. The Figures 4.1 and 4.2 depicts the proposed method. Figure 4.1 shows that when clustering is applied on the weekly average usage data of customers, the customers are grouped in different clusters. Figure 4.2 shows that once the clustering is done, training instances related to the members of a cluster is combined to train classifier for that cluster. At runtime, a customer will be grouped in a cluster based on its weekly usage. Once the program knows the cluster assigned to a customer, the program will load the corresponding classifier to make electricity demand forecasts about the customers.

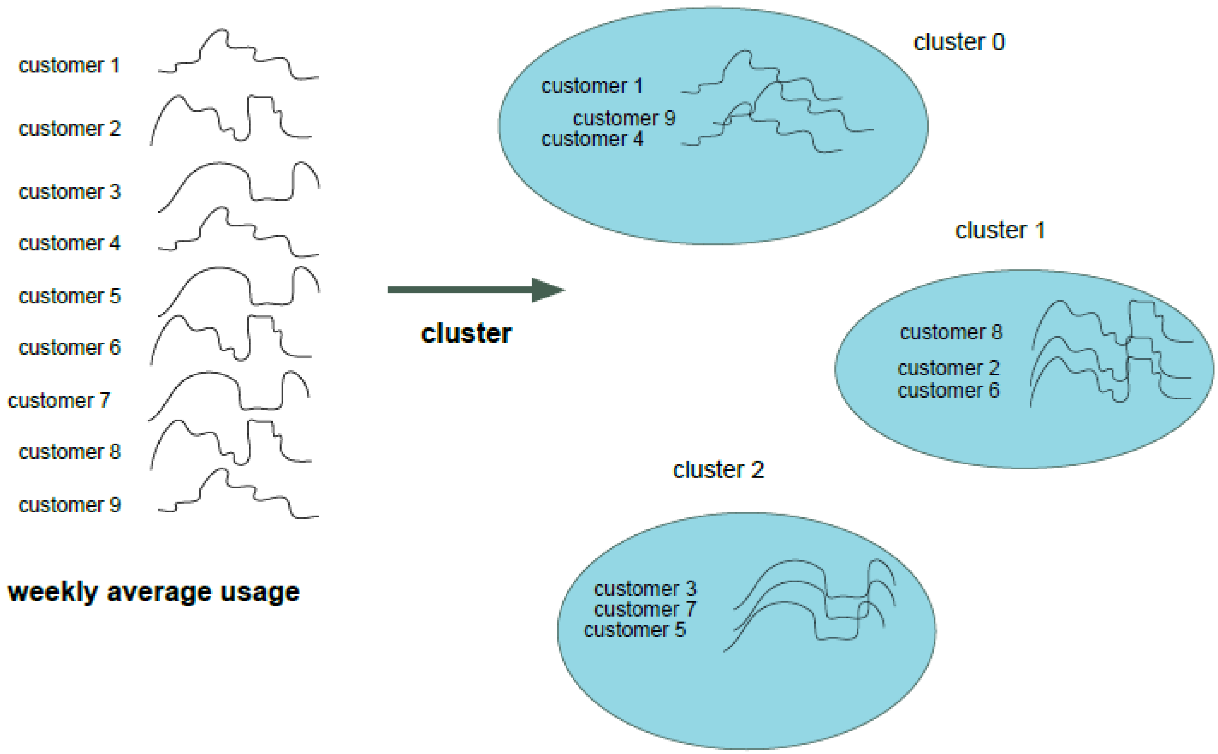

Figure 4.1: Clustering customers based on weekly average usage.

Algorithm 7 shows the procedure for extracting average electricity usage of each hour of a week. Next, all the average weekly usages are combined together to make training set for the clustering algorithm. I have used k-means [37] clustering algorithm to cluster the training set. I have trained clusters of sizes 1 to 18. Algorithm 8 describes the procedure for making clusters from the training instances. Once a k-means of cluster size k is made, a program groups the hourly usages of the customers in the same cluster and combines them to make training set for machine learning classifier. This training set is used to train a linear regression classifier. Algorithm 9 describes how the cluster-based predictor's performance was evaluated. I fixed the optimal number of clusters from two observations. First, I used Expectation Maximizaion (EM) algorithm to see how many clusters might be present in the data. If number of clusters are unknown, EM algorithm can tell how many clusters are

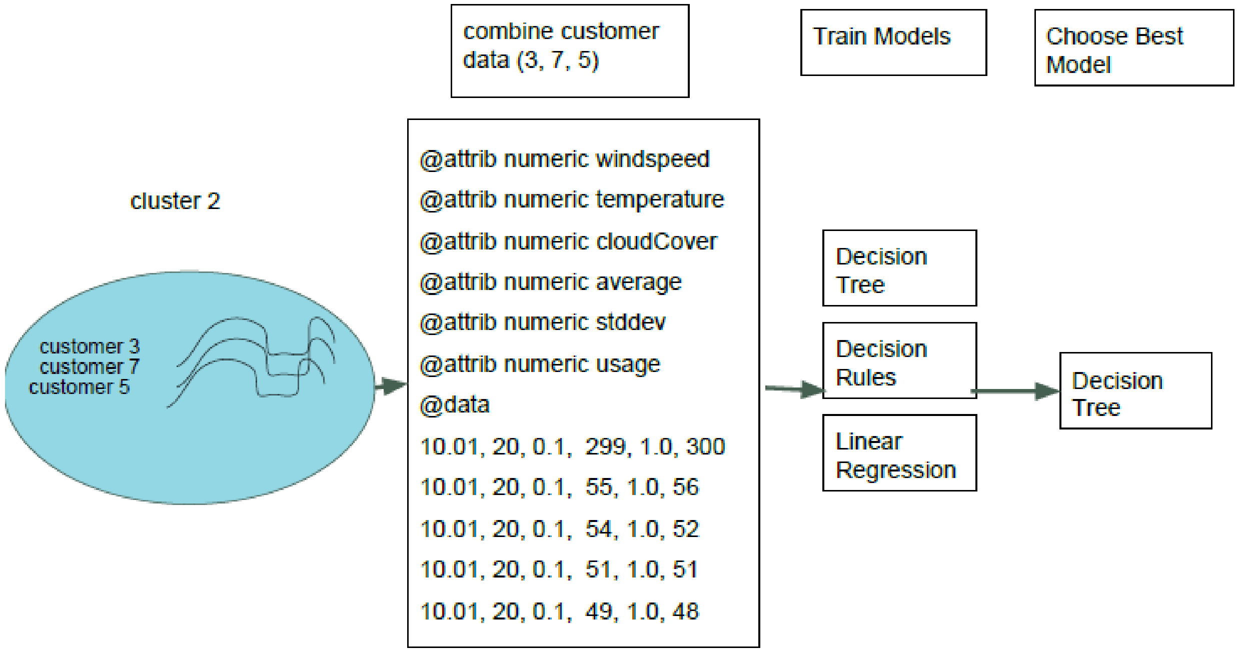

Figure 4.2: Training prediction classifier for a cluster.

there. Secondly, I noticed the prediction error produced for each cluster of size k in kmeans clustering algorithm and observed for what size of cluster the error is minimum. The result is described in Section 4.9.1. Once the number of the clusters was fixed, a program creates several machine learning predictors to see which one performs best for a given cluster. The machine learning classifiers that were tried out are Linear Regression, M5P Rules, M5 Tree, REP tree[37].

4.7 Testing Performance

Algorithm 11 shows how the accuracy of each prediction model was tested. Each instance of the test set was tested against each demand prediction scheme. The Mean Absolute Percentage Error (MAPE) was calculated for each scheme [37]. Algorithm 12 shows the algorithm for computing error.

Algorithm 7 write average electricity usage of the customers of each hour of the week

Require: information of all timeslots has been received

```
for each customer do
    trainingInstance = create empty training instance
    for each day of week do
        for each hour of day do
            averageUsage = get average usage of day and hour of customer
            append averageUsage to the trainingInstance
        end for
    end for
    writeToAvgUsageFile(trainingInstance)
end for
```

Algorithm 8 create kmeans cluster of size k from weekly usage training instance file

```
data = load weekly average usage file
kmeansCluster = build kmeans cluster of size k based on data
save kmeansCluster
```

4.8 Experimental Results

The following subsections describe the results at each stage of the experiments. The stages are 1) finding optimal number of clusters. 2) Once the number of clusters has been fixed, a different classifier needs to be made for each cluster to see which one makes the best forecast for a particular cluster. 3) After that, the baseline predictor that needs a classifier for each customer needs to be trained. At this point, for each customer several classifier has been tried out to see which classifier makes best demand forecast about that customer. 4) Finally, the proposed mechanism is tested against the two baseline demand forecasting methods.

Algorithm 9 find error of kmeans clusters of different size

```
for each cluster size k do
    get the kMeansCluster of size k
    for cluster in KMeansCluster do
        combine slot based training instances of that cluster
        train linear regression classifier based on the combined data
        save the classifier for cluster
    end for
end for
for each training instance do
    compute error of the instance using each kMeansCluster
end for
```

Algorithm 10 find best classifiers of each cluster of kmeans cluster of size k

```
for each cluster in kMeansCluster do
    combine slot based data of the all the customers in cluster
    train available classifiers on the combined data using 10 fold cross validation
    choose the classifier with minimum error
    save the classifier for making demand forecasting for cluster
end for
```

4.9 Finding Individual Predictor for Each Customer

Based on the data from each of the customers, the four types of classifiers described previously were tested. For each customer, the Table 4.1 shows the best performing classifier for each customer.

The Figure 4.3 shows error percentage of each of the predictors type for each of the customer types.

Algorithm 11 performance evaluation of each method

for each test instance **do**
    classify the test instance using moving average usage [**algorithm 2**]
    classify the test instance using individual prediction mechanism
    classify the test instance using cluster based predictor
    calculate and accumulate errors of each mechanism [**algorithm12**]
    update moving average baseline predictor based on the information from the test instance [**algorithm 1**]
end for
find average error from the accumulated errors for each forecasting mechanism

Algorithm 12 calculate error from the predicted value and the true value

absoluteError = abs(predictedValue - trueValue)
relativeAbsoluteError = (absoluteError / trueValue) * 100 %

4.9.1 Finding the optimal number of clusters

First, I segmented the customer using KMeans clustering algorithm with cluster sizes from 1 to 18. I choose the number 18 because there are 17 consumption type customers in Power TAC. In worst case, each customer may be different and there will be 17 clusters. For kMeans Euclidian Distance [37] was used for similarity metric, max number of iteration [37] was set to 500. For KMeans with size k, we will have k clusters. For each of the k clusters, I had a linear regression predictor. I observed the relative percentage error and absolute average the above cluster sizes. Figure 4.4 shows the cluster size vs MAPE. It shows that after cluster of size 4, the size of the cluster does not have a big impact on the prediction performance. The EM algorithm reports that there are 4 clusters present in the data when max iteration parameter [37] was set to 100, minimum standard deviation parameter [37] was set to 1.0E-06 and num clusters parameter [37] was set to -1. To keep things simple, I have decided to choose Kmeans cluster of size 4. When k = 4 was chosen, the Table 4.2 shows the cluster assignment for each customer. The Table 4.2 shows that,

Table 4.1: Best individual predictor for each customer

Customer Name	Best Predictor Type
BrooksideHomes	M5P
CentervilleHomes	M5P
DowntownOffices	M5P
EastsideOffices	M5P
OfficeComplex 1 NS Base	LinearRegression
OfficeComplex 1 SS Base	LinearRegression
OfficeComplex 2 NS Base	LinearRegression
OfficeComplex 2 SS Base	LinearRegression
Village 1 NS Base	M5P
Village 1 RaS Base	LinearRegression
Village 1 ReS Base	M5P
Village 1 SS Base	M5P
Village 2 NS Base	LinearRegression
Village 2 RaS Base	M5P
Village 2 ReS Base	M5P
Village 2 SS Base	M5P
MedicalCenter@1	M5P

cluster 0 held most of the offices, cluster 2 held most of the village types, cluster 3 held the medical center, cluster 1 held large housing such as brooksidehomes, centerville homes and large offices such as downtown offices and centerville offices. This observation suggests that the office, village, large corporation customers have similar kind of electricity usage pattern in Power TAC.

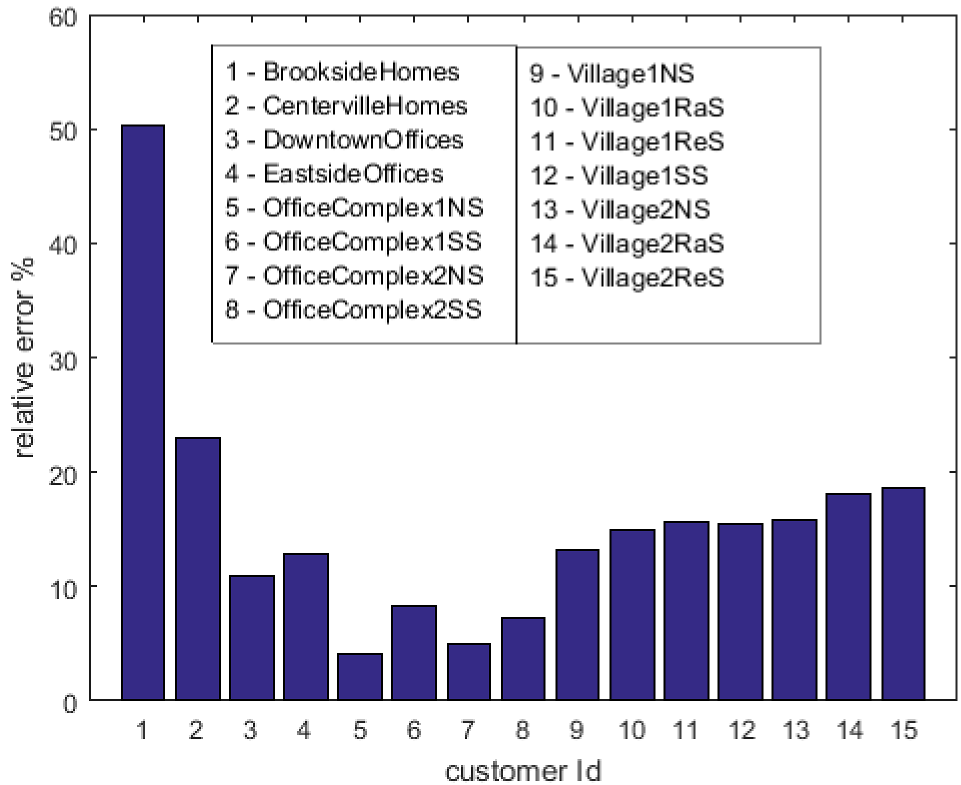

Figure 4.3: Performance of the best classifier for each customer type. Customer Medical center was excluded as it was showing huge error.

4.9.2 Finding the best predictor for each cluster

Once the features are extracted, I have tested M5Tree, Linear Regression, M5P rules and REP tree machine learning classifiers to see which one performs the best for each of the 4 clusters. Figure 4.5, 4.6, 4.7, 4.8 show the average relative percentage errors that each classifier produced for cluster 0, 1, 2, and 3 respectively. Figure 4.5 shows that among all the classifiers, M5P produces the minimum amount of error. So M5P will be used as the demand predictor for cluster 0. For cluster 1, 2 and 3 the M5P, REPTREE and M5RULES will be used respectively as demand predictors.

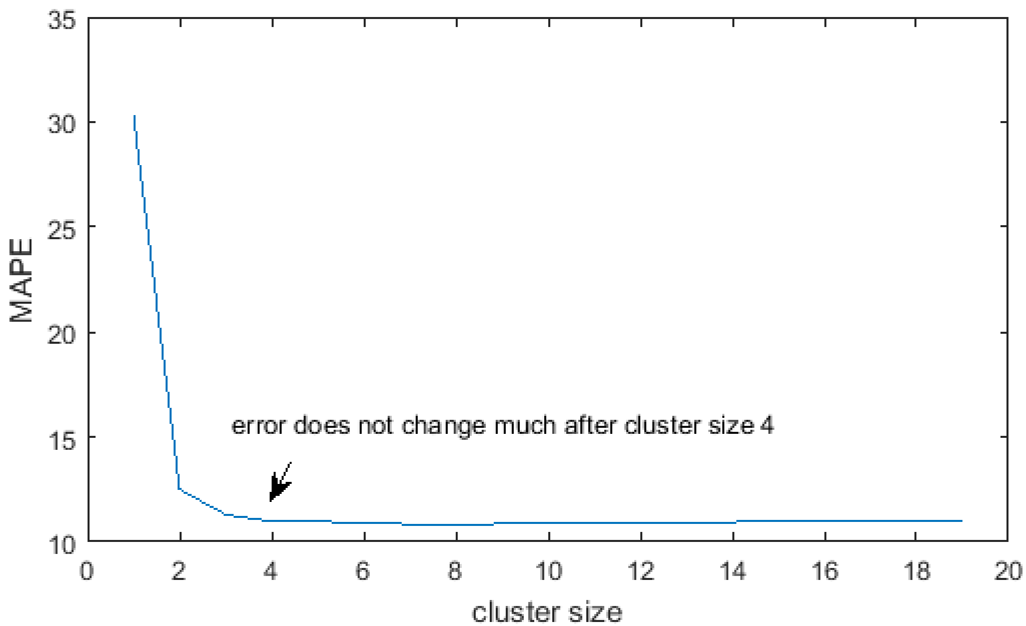

Figure 4.4: Kmeans cluster size vs MAPE.

4.9.3 Comparison Among the Three Prediction Schemes

Finally, all three demand prediction schemes were tested with test set. From Figure 4.9, we can see that cluster based prediction mechanism performed almost as good as the mechanism individual prediction scheme, and it did much better than the default moving average prediction scheme. Individual prediction mechanism can be considered as a cluster based prediction mechanism with cluster size n. The cluster based mechanism used only 4 clusters, yet its performance was almost as same as the individual prediction mechanism. This suggests the importance of the cluster based prediction mechanism; it will need less amount of memory, the system is simple and generalized yet it produces quality prediction.

4.9.4 Model Accuracy and Training Set Size

Figure 4.10 shows the error of all three models with increasing training set size. It appears that the accuracy of the models does not change much with increased training set size.

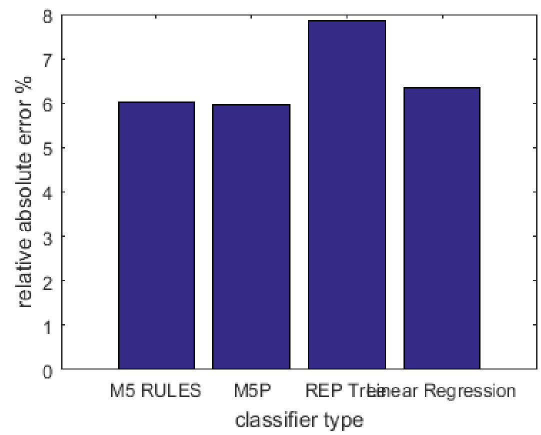

Figure 4.5: Cluster 0

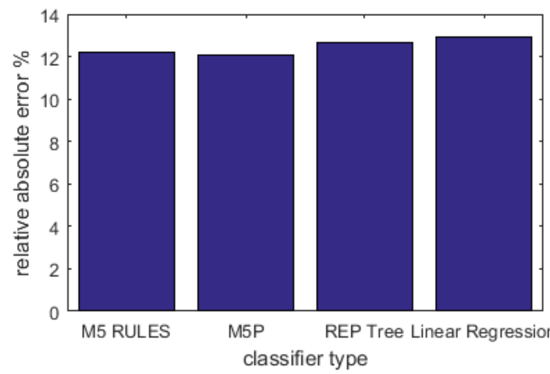

Figure 4.6: Cluster 1

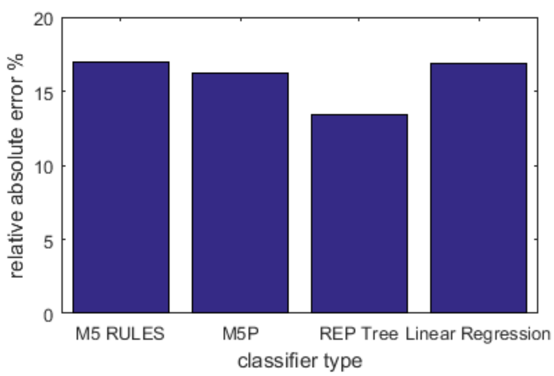

Figure 4.7: Cluster 2

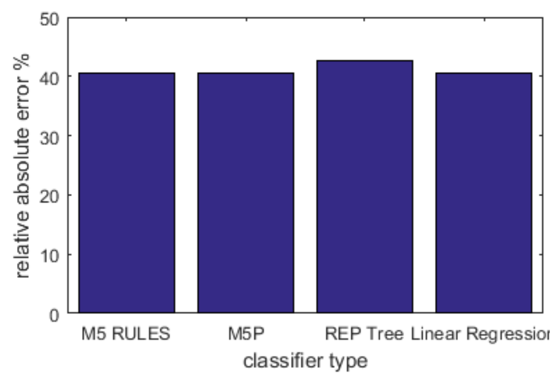

Figure 4.8: Cluster 3

Table 4.2: Assigned cluster for each customer

Customer Name	Assigned Cluster Number
BrooksideHomes	0
CentervilleHomes	0
DowntownOffices	1
EastsideOffices	1
OfficeComplex 1 NS Base	0
OfficeComplex 1 SS Base	0
OfficeComplex 2 NS Base	0
OfficeComplex 2 SS Base	0
Village 1 NS Base	2
Village 1 RaS Base	2
Village 1 ReS Base	2
Village 1 SS Base	2
Village 2 NS Base	2
Village 2 RaS Base	2
Village 2 ReS Base	2
Village 2 SS Base	2
MedicalCenter@1	3

The reason behind this reason may be because the models were trained based on past tournament data. During a tournament the customer behavior remains unchanged, so adding more data may be redundant.

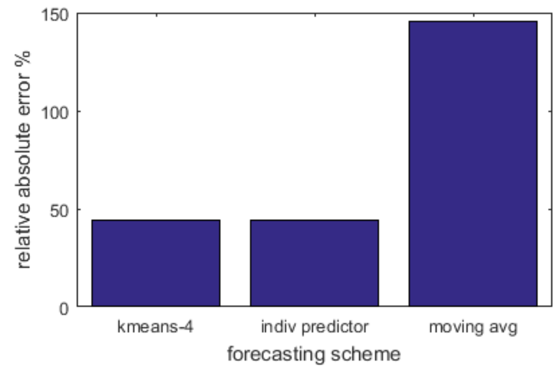

Figure 4.9:
Comparison among three prediction schemes

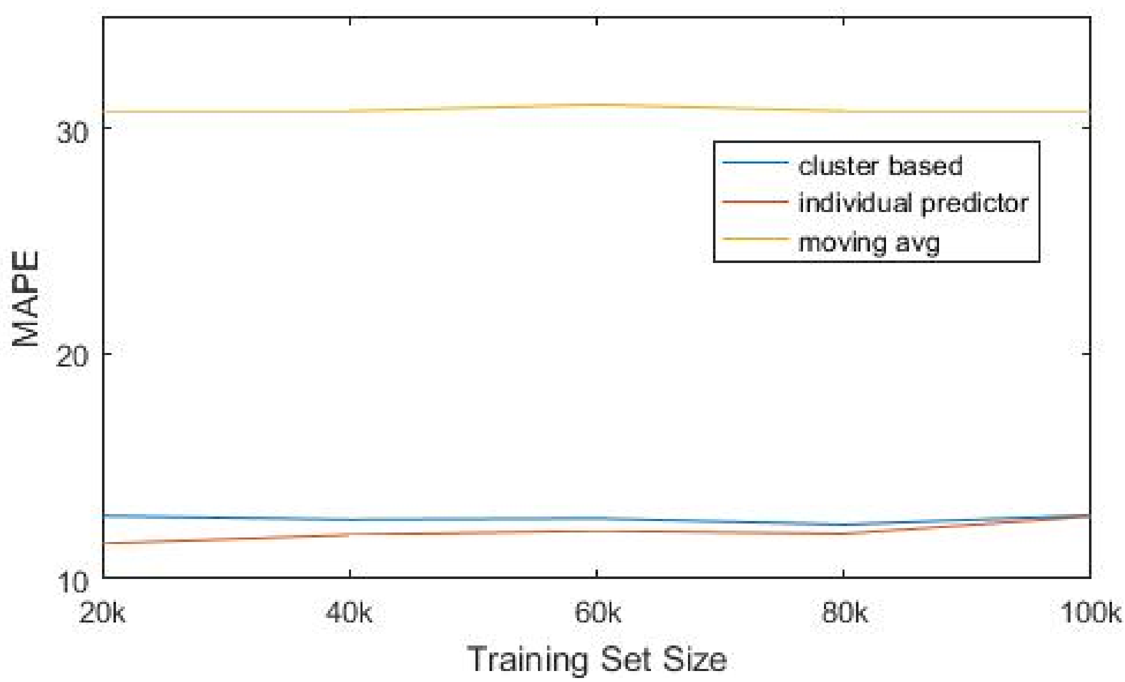

Figure 4.10: Comparison among three prediction schemes

Chapter 5

Conclusions and Future Work

My work demonstrates that a single demand predictor may not be suitable for the Power TAC scenario. It was clear that for each power type, a broker should use different demand forecasting mechanism. For consumption type customers a small set of features containing temperature, cloud cover, wind speed, average electricity usage and standard deviation of electricity usage produced good results. I provided evidence that cluster based demand prediction mechanism was a highly dependable demand prediction scheme. I also showed that for consumption type customers in the Power TAC simulation, the size of the cluster does not matter when size of the cluster is greater than or equal 4. The individual predictor scheme can be considered as a forecasting mechanism with 17 clusters. The proposed forecasting methodology was able to achieve very similar demand forecasting accuracy using only 4 clusters instead of 17 clusters. The individual predictor scheme has a serious weakness, in fact the demand predictors are hardcoded by the customer names. If during the simulation the name of a customer is changed, this mechanism will not work. On the other hand, the proposed mechanism can be trained on previous game logs and does not have the problem with handling new customers. Finally, I showed the accuracy of the model may not depend on the size of the training set if customer properties are not changed in server.

5.1 Future Work

My work so far only deals with demand forecasting for the consumption customers. I would like to compare the proposed method against other baselines. Instead of using simulation data, real-world data can be used to test the effectiveness of the proposed method. It may be the case that there are more helpful features out there, I would like to do more experiment with new features. Instead of using simple classifiers, I would like to train complex classifiers such as Neural Networks [37] and Support Vector Machines [37] to see if the accuracy can be improved more. The proposed mechanism seems to be applicable for solar energy production customers with a slight change. For customer with irregular demand pattern such as customers with demand shifting capabilities and the electric vehicle customers, different technique of demand forecasting has to be figured out.

References

[1] Matlab. http://www.iea.org/statistics/.

[2] HM Al-Hamadi and SA Soliman. Short-term electric load forecasting based on kalman filtering algorithm with moving window weather and load model. *Electric power systems research*, 68(1):47–59, 2004.

[3] Nima Amjady. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems*, 16(3):498–505, 2001.

[4] G Andersson, P Donalek, R Farmer, N Hatziargyriou, I Kamwa, P Kundur, N Martins, J Paserba, P Pourbeik, J Sanchez-Gasca, et al. Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance. *IEEE transactions on Power Systems*, 20(4):1922–1928, 2005.

[5] Bo-Juen Chen, Ming-Wei Chang, and Chih-Jen Lin. Load forecasting using support vector machines: A study on eunite competition 2001. *Power Systems, IEEE Transactions on*, 19(4):1821–1830, 2004.

[6] MY Cho, JC Hwang, and CS Chen. Customer short term load forecasting by using arima transfer function model. In *Energy Management and Power Delivery, 1995. Proceedings of EMPD'95., 1995 International Conference on*, volume 1, pages 317–322. IEEE, 1995.

[7] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart gridthe new and improved power grid: A survey. *Communications Surveys & Tutorials, IEEE*, 14(4):944–980, 2012.

[8] Heiko Hahn, Silja Meyer-Nieberg, and Stefan Pickl. Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research*, 199(3):902–907, 2009.

[9] Melissa Hart and Richard de Dear. Weather sensitivity in household appliance energy end-use. *Energy and Buildings*, 36(2):161–174, 2004.

[10] Luis Hernández, Carlos Baladrón, Javier M Aguiar, Belén Carro, and Antonio Sánchez-Esguevillas. Classification and clustering of electricity demand patterns in industrial parks. *Energies*, 5(12):5215–5228, 2012.

[11] Ku-Long Ho, Yuan-Yih Hsu, Chuan-Fu Chen, Tzong-En Lee, Chih-Chien Liang, Tsau-Shin Lai, and Kung-Keng Chen. Short term load forecasting of taiwan power system using a knowledge-based expert system. *IEEE Transactions on Power Systems*, 5(4):1214–1221, 1990.

[12] Ching-Lai Hor, Simon J Watson, and Shanti Majithia. Analyzing the impact of weather variables on monthly electricity demand. *IEEE transactions on power systems*, 20(4):2078–2085, 2005.

[13] Che-Chiang Hsu and Chia-Yon Chen. Regional load forecasting in taiwan—-applications of artificial neural networks. *Energy conversion and Management*, 44(12):1941–1949, 2003.

[14] Ercan Izgi, Ahmet Öztopal, Bihter Yerli, Mustafa Kemal Kaymak, and Ahmet Duran Şahin. Short–mid-term solar power prediction by using artificial neural networks. *Solar Energy*, 86(2):725–733, 2012.

[15] Nur Adilah Abd Jalil, Maizah Hura Ahmad, and Norizan Mohamed. Electricity load demand forecasting using exponential smoothing methods. *World Applied Sciences Journal*, 22(11):1540–1543, 2013.

[16] Wolfgang Ketter, John Collins, and Prashant Reddy. Power tac: A competitive economic simulation of the smart grid. *Energy Economics*, 39:262–270, 2013.

[17] Wolfgang Ketter, John Collins, and Mathijs De Weerdt. The 2016 power trading agent competition. *ERIM Report Series Reference*, 2016.

[18] Bart Liefers, Jasper Hoogland, and Han La Poutré. A successful broker agent for power tac. In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*, pages 99–113. Springer, 2014.

[19] Rab Nawaz Lodhi and RK Malik. Impact of electricity shortage on daily routines: A case study of pakistan. *Energy & Environment*, 24(5):701–709, 2013.

[20] Paras Mandal, Tomonobu Senjyu, and Toshihisa Funabashi. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy Conversion and Management*, 47(15):2128–2142, 2006.

[21] Fintan McLoughlin, Aidan Duffy, and Michael Conlon. A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied energy*, 141:190–199, 2015.

[22] Ibrahim Moghram and Saifur Rahman. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Transactions on power systems*, 4(4):1484–1491, 1989.

[23] S Ozdemir and R Unland. Agentude: The success story of the power tac 2014s champion. In *AAMAS Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis (AMEC/TADA 2015)*, 2015.

[24] Serkan Ozdemir and Rainer Unland. A winner agent in a smart grid simulation platform. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, volume 2, pages 206–213. IEEE, 2015.

[25] Jaime Parra Jr and Christopher Kiekintveld. Initial exploration of machine learning to predict customer demand in an energy market simulation. In *Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

[26] Cameron W Potter, Allison Archambault, and Kenneth Westrick. Building a smarter smart grid through better renewable energy information. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–5. IEEE, 2009.

[27] Hao Quan, Dipti Srinivasan, and Abbas Khosravi. Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE transactions on neural networks and learning systems*, 25(2):303–315, 2014.

[28] Saifur Rahman and Rahul Bhatnagar. An expert system based algorithm for short term load forecast. *Power Systems, IEEE Transactions on*, 3(2):392–399, 1988.

[29] Andre Richter, Erwin van der Laan, Wolfgang Ketter, and Konstantina Valogianni. Transitioning from the traditional to the smart grid: Lessons learned from closed-loop supply chains. In *Smart Grid Technology, Economics and Policies (SG-TEP), 2012 International Conference on*, pages 1–7. IEEE, 2012.

[30] Ina Rüdenauer and Carl-Otto Gensch. Energy demand of tumble driers with respect to differences in technology and ambient conditions. *Öko-institut, Freiburg*, 2004.

[31] Arun P Sanghvi. Economic costs of electricity supply interruptions: Us and foreign experience. *Energy Economics*, 4(3):180–198, 1982.

[32] Nicholas I Sapankevych and Ravi Sankar. Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine*, 4(2):24–38, 2009.

[33] Andrea Schreiber, Petra Zapp, and Wilhelm Kuckshinrichs. Environmental assessment of german electricity generation from coal-fired power plants with amine-based carbon capture. *The International Journal of Life Cycle Assessment*, 14(6):547–559, 2009.

[34] Jonathan Serrano, Enrique Munoz de Cote, and Ansel Y Rodríguez. Fixing energy tariff prices through reinforcement learning.

[35] Daniel Urieli and Peter Stone. Tactex'13: a champion adaptive power trading agent. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1447–1448. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[36] Xishun Wang, Minjie Zhang, Fenghui Ren, and Takayuki Ito. Gongbroker: A broker model for power trading in smart grid markets. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, volume 2, pages 21–24. IEEE, 2015.

[37] Ian H Witten and Eibe Frank. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann, 2005.

Curriculum Vitae

Saiful Abu grew up in a small city called Khulna in Bangladesh which is surrounded by rivers and forests. He is a first generation college graduate coming of a family with aquaculture and aquaculture background. He went to the most prestigious engineering school in Bangladesh University of Engineering and Technology to study Bachelor of Science in Computer Science. He graduated from there in 2012. After working for a few years in the industry, he came to the University of Texas at El Paso in Fall 2014. There he worked as a Teaching Assistant under the supervision of Dr. Olac Fuentes and Dr. Julio Urenda. He also worked as a Research Assistant under the supervision of Dr. Christopher Kiekintveld. He graduated with a Master of Science degree in Computer Science in Summer 2016.

Permanent address: 52, Jahidur Rahman Sarak
Khulna, Bangladesh.