A Study on Various Methods for Mining Energy Consumption pattern from Smart Home Big Data

Jemeema Maria John
P.G Scholar, Department of Computer Science and Engineering, St Joseph’s College of Engineering and Technology, Palai, India
E-mail: jemeemajohn@gmail.com

Hima Anns Roy
Assistant Professor, Department of Computer Science and Engineering, St Joseph’s College of Engineering and Technology, Palai, India
E-mail: himaannsroy@gmail.com

Abstract. Houses are the main source of electrical energy consumption. Getting information about energy consumption pattern from smart homes are of vital importance. It helps in analyzing the behaviour of inhabitants and that of the equipments. Smart meters in homes collect large volume of data each day. Data mining techniques can be used to get relevant information from the smart home big data. Data mining helps in easily predicting the consumption patterns of various equipments. Forecasting energy consumption pattern helps in better energy management plans. The techniques that are used for mining and predicting energy consumption patterns are mentioned in this paper.

Keywords: Smart homes, Smart meter, big data, energy consumption, data mining.

1. Introduction
In todays world electrical energy is a major source of power generation [1]. The main consumers of electricity are houses. The major consumers of electricity are the various appliances that are used. Knowledge about the energy consumption pattern helps people in various ways. It helps in analyzing the consumption of electricity by appliances and devices. The prediction of energy consumption pattern or appliance usage pattern is also of greater importance. Energy consumption across the world is rapidly increasing due to various reasons. Higher energy consumption rates can occur due to various reasons. Some of them are increase in population, pressures for better living standards, large-scale industrialization in developing countries. Thus, for accurate investment planning of energy production generation and distribution a forecasting technique is essential.

It is very difficult to analyze millions of data records produced by smart meters. There are a lot of techniques available for analyzing smart meter big data. But these techniques are not very efficient. So a new technique called data mining was adopted. Data mining techniques can be used to extract relevant information from large amount of data called big data. The data mining techniques play a big role in obtaining energy consumption patterns from smart meter.
2. Data Mining

Data mining is the process of extracting relevant information from large collection of data. It can also be defined as analysing large data sets to find patterns to establish relationship among various factors. Data mining can be applied to extremely large collection of data sets. Generally data mining can be applied to any kind of data as long as the data are meaningful for a target application. The most basic forms of data for mining applications are database data, data warehouse data, and transactional data.

Data mining processes are very useful in collecting required information from big data. Big data represents huge volume of data. Big data sets are those that outgrow the simple kind of database and are not very easy to analyse, but with data mining techniques analysis on these big data sets are made easy. Thus datamining techniques are used for analyzing big data.

Some healthcare applications utilize data mining techniques to find out activity patterns of humans in smart homes. This information can be used to understand the life style of human and to identify any deviation from their daily routines. This is done by mainly analysing the appliance usage, appliance to appliance and appliance to appliance to time relationships [2].

3. Mining and Predicting Energy Consumption Patterns

3.1. DCMiner Algorithm

Based on the usage representation, Dynamic Correlation Miner(DCMiner) [3] algorithm is proposed to incrementally discover correlation patterns. DCMiner can incrementally mine the correlation patterns from evolving usage databases effectively and efficiently.

3.2. CTMiner Algorithm

Mining temporal patterns from time interval-based data is a difficult problem since the processing of complex relations among intervals may require generating and examining large amount of intermediate subsequence. A novel technique, called incision strategy and a new representation, known as coincidence representation are used to solve this problem.

An algorithm called CTMiner(Coincidence Temporal Miner) [4] is developed to discover frequent time-interval based patterns. The algorithm uses two pruning techniques to reduce the search space effectively. The pruning techniques used are called pre pruning and post pruning.

3.3. Incremental Mining

The key elements for understanding power consumption of a typical home are related to the activities the inhabitants are performing, the time at which appliances are used, and the inter-dependencies among appliances. This information can be obtained from context rich smart meters big data.

Furthermore, generation of energy consumption data from a smart meter is an ongoing continuous process and over period of time inter-appliance associations can change or new ones can establish. An incremental model for mining frequent power consumption pattern [5] is proposed. Here frequent patterns are extracted from data comprising of appliance usage tuples for 24 hours period, in a progressive manner.

3.4. Progressive Incremental Mining

Since houses are equipped with smart meters, appliance usage can be easily collected and stored in database. Analysing the database will help in studying the energy consumption behavioural pattern. Appliance-appliance and appliance-time relation are determined for getting more details.

Due to the continuous generation of energy consumption data, over a period of time, appliance associations can change. Therefore, they need to be captured regularly and continuously.
Progressive incremental data mining mechanism can be applied to energy consumption data to overcome these challenges [6].

Progressive incremental data mining mechanism can be done by four distinct phases: data preparation, frequent pattern mining and clustering, association rules generation, and visualization. Using the progressive mining techniques AoI (Appliances of Interest), the major energy consuming appliances for the house, which require equal attention from both consumers and utilities to attain energy efficiency.

3.5. Decision tree and time based multi label classifier
Buildings represent the first energy consumer. Home automation system is made up of household appliances that are linked via a communication network which allows interactions for control purposes. Load management helps inhabitants to adjust their power consumption. The energy management of residential houses requires the prediction of future usage of appliances.

This model tries to formalize such an approach using a time-series based multi-label classifier [7]. The classifier takes into account correlation between different appliances. Here, prediction results are shown for 1-hour.

The benefit of using only energy consumption reading for each appliance is many-fold. The objective of this model is to construct a system that is able to predict the appliance usage in housing which helps the system to organize energy production and consumption and to decide which appliance will be used at each hour.

3.6. Rule mining and DBSCAN clustering
Energy management in home is a key factor for creating a sustainable future for our society. Now a days a lot of tools are available for measuring energy usage in homes. One such model uses a technique called rule mining [8]. The model uses a rule mining algorithm to identify significant associations between energy usage and four key features. These key features are hour of the day, day of the week, use of other appliances in the home, and user-supplied annotations of activities such as working or cooking.

This model also utilizes the DBSCAN clustering method for grouping the various activities. DBSCAN is used to produce a unique profile for each device in the home. DBSCAN is a density-based algorithm that identifies clusters while excluding noise. DBSCAN is very helpful for as it does not require the number of clusters to be provided as input and it can be implemented very efficiently.

3.7. Clustering and Bootstrapping
Large quantity of information about, how customers use their energy is available through smart meters. Clustering the most important attributes of customers is a very common method for better understanding the residential energy behaviours that exist and has many applications [9].

Clustering methods can be applied to smart meter data. It helps distribution network operators (DNOs) to manage and plan the low voltage networks. This method proposes an in-depth analysis of customer smart meter data. Here four key time periods are used in which the data should be analysed. The time periods used are overnight, breakfast, daytime and evening.

The time periods not only helps to model peak time behaviours and their variation but also allows to reduce the number of attributes in the clustering implementation. It uses 10 distinct behaviour groups that describes the customers based on their demands. Finally, a bootstrapping technique is used to make the clustering reliable.
| Techniques                        | Pros                                                                 | Cons                                                                 |
|----------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| DCMiner Algorithm                | • Takes less memory space for execution.                           | • As the size of the database increases the processing time of DCMiner increases |
| CTMiner Algorithm                | • Reduces search space                                              | • Requires additional computations.                                   |
|                                   | • Consumes much smaller memory space.                               |                                                                      |
| Incremental mining                | • Process only the incremental part, so execution time is faster.   | • Costly in case of small amount of data.                             |
| Progressive incremental mining    | • Execution time is faster.                                         | • Costly in case of small amount of data.                             |
| Decision tree and time based multi label classifier | • It takes into account both conditional and unconditional label dependencies. | • Process of labeling to multi-label data is expensive and time consuming |
| Rule mining and DBSCAN clustering | • It does not require the number of clusters to be provided as input | • It does not work well in case of data having high dimensionality    |
|                                   | • can be implemented very efficiently                              |                                                                      |
| Clustering and bootstrapping      | • Allows an experimenter to adjust results to reflect biases in their sample. | • The success depends on the data set it is used on.                  |
|                                   | • Helps to create more accurate customer profiles.                  | • Small differences in input can lead to substantial differences in output. |
4. Conclusion
In this study, various methods that were used to analyse energy consumption pattern were discussed. It can be seen that a lot of techniques are used for obtaining the energy patterns. Data mining helps in extracting the energy patterns from smart home big data. It also helps in forecasting the consumption patterns of energy usage. As the world population is increasing day by day energy consumption is also increasing. So analyzing the energy consumption patterns helps us to understand the user behaviour and to optimize the power consumption accordingly. Using the monthly or daily data can help make accurate predictions. For example, by analyzing on more amount data, predictions can be done with less error rate. Since data mining plays huge role in healthcare sector, the forecasted results can be used in healthcare sector for predicting human behavioural anomalies.

5. References
[1] Noorollah Karimtabar, Sadegh Pasban, Siavash Alipour, Analysis and predicting electricity energy consumption using data mining techniques, A case study I.R. Iran Mazandaran province, 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA 2015) March 11-12, 2015.
[2] Abdulsalam Yassine; Shailendra Singh; Atif Alamri, Mining Human Activity Patterns From Smart Home Big Data for Health Care Applications, IEEE 2017.
[3] Andreas B. Y. Chiang, H.C. Hung, P. J. Chen, S. Y. Peng, Y. C. Chen, Incrementally mining usage correlations among appliances in smart homes, in Proc. 18th Int. Conf. Netw.-Based Inf. Syst. (NBiS), pp. 273-279, 2015.
[4] J. Jiang, S. Lee, W. Peng, Y. Chen, An Efficient Algorithm for Mining Time Interval-based Patterns in Large Databases, Proceedings of 19th ACM International Conference on Information and Knowledge Management (CIKM10), pp. 4958, 2010.
[5] A. Yassine, S. Shirmohammadi, S. Singh, Incremental mining of frequent power consumption patterns from smart meters big data, in Proc. IEEE Electr. Power Energy Conf. (EPEC), pp. 16, Oct. 2016.
[6] A. Yassine, S. Singh, Mining energy consumption behaviour patterns for households in smart grid, IEEE Trans. Emerg. Topics Comput., to be published...
[7] K. Basu, S. Bacha, V. Debusschere, Appliance usage prediction using a time series based classification approach, in Proc. IEEE 38th Annu. Conf. Ind. Electron. Soc. (IECON), pp. 1217-1222, May 2012.
[8] N. Banerjee, S. Rollins, Using rule mining to understand appliance energy consumption patterns, in Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference, pp. 2937-3 2014.
[9] C. Singleton, P. Grindrod, S. Haben, Analysis and clustering of residential customers energy behavioral demand using smart meter data, IEEE Transactions on Smart Grid, vol. 7, no. 1, pp. 136144, 1 2016.