Big Data Mining of Smart Home Appliance Energy Time Series for Behavioral Analytics and Human Activity Prediction.

Ms.Sara Sarwath, Ms.Afreen Sultana

Muffakhamjah College of Engineering & Technology

ABSTRACT-Nowadays, there's an ever-increasing migration of individuals to urban areas. Health care services are one among the foremost difficult aspects that's greatly suffering from the huge flow of individuals to town centers. Hence, almost all cities across the globe are investing heavily in digital transformation so as to produce healthier ecosystem for individuals. In such transformation, several homes are being equipped with smart devices (e.g. intelligent meters, sensors etc.) that generate large volumes of fine-grained and factual information that may be analyzed to support smart town services. In this paper, we tend to propose a model that utilizes smart home big data as a way of learning and discovering human activity patterns for health care applications. We tend to propose the utilization of frequent pattern mining, cluster analysis and prediction to determine and analyze energy usage changes sparked by occupants' behavior. Since people's habits are largely known by their everyday routines, discovering these routines permits us to acknowledge abnormal activities that will indicate people's difficulties in taking care of themselves, corresponding to not getting food ready or not taking shower/bath. Our work addresses the necessity to explore temporal energy consumption patterns at the appliance level, as it directly involves human activities. For the assessment of the proposed mechanism, the analysis uses the UK Domestic Appliance Level Electricity data set (UK-Dale) - statistical data of power consumption collected from 2012 to 2015 with time resolution of six seconds for 5 homes with 109 appliances from Southern England. The information from smart meters is recursively mined within the quantum/data slice of twenty four hours, and also the results are maintained across consecutive mining exercises. The Hadoop ecosystem is used to process the big data generated from the smart homes and thus enabling the distributed processing. The results of distinctive human activity patterns from appliance usage are mentioned in detail along with accuracy of short and long run predictions.

KEYWORDS—*Big data, Cluster Analysis, Frequent Pattern, Hadoop, Health care, Incremental Data-Mining, Prediction, Smart homes.*

1. INTRODUCTION

Studies demonstrate that by year 2050, 66% of the total populace will live in urban regions. The demand for healthcare resources will be extraordinarily influenced by this immense convergence of individuals to downtown areas. This phenomenal statistic change places enormous burden on urban communities to reevaluate the conventional methodologies of giving wellbeing administrations to inhabitants. In reacting to the new needs and difficulties, urban communities are right now grasping huge digital transformation with an end goal to help feasible urban groups, also, give healthier environment.

In such change, a huge number of homes are being furnished with smart devices (e.g. smart meters, sensors and so on.) which produce huge volumes of fine-grained and indexical information that can be broke down to help support healthcare services. Headway of big data mining technologies, which give methods for preparing huge measure of information for significant bits of knowledge, can help us in seeing how individuals approach their life. For instance, observing the progressions of appliance utilization inside a smart home can be utilized to in a roundabout way decide the individual's prosperity in view of historical data. Since individuals' propensities are generally distinguished by regular schedules, finding these schedules enables us to perceive peculiar exercises that may demonstrate individuals' challenges in taking administer to themselves, for example, not getting ready nourishment or not using shower/shower. The fundamental connection between's machine use inside the smart home and routine exercises can be utilized by health applications to distinguish potential medical issues. This isn't just going to ease the load on human services frameworks, yet additionally giving 24 hour checking administration that naturally recognize typical and anomalous practices for autonomously living patients or those with self-constraining conditions (e.g. elderly and patients with intellectual disabilities).

We propose the utilization of vitality information from smart meters introduced at homes to divulge critical exercises of occupants. Disaggregated power utilization readings are specifically related to the exercises performed at home. For example, if the "Stove" is ON, the operation of this machine is in all probability related with action "Getting ready Food". The time (e.g. morning or night) of this operation may likewise show the sort of the feast, for example, breakfast or supper.

Moreover, individuals frequently perform more than one action at the same time, for example, "Getting ready Food"

and "Tuning in to Music" or on the other hand "Staring at the TV", which implies different machines are worked together. In this unique situation, we investigate shoppers' worldly vitality utilization designs at the appliance level to recognize different appliance uses and anticipate their operations over short and long haul time allotments. This is especially conceivable without extra equipment since the smart meter information has time-arrangement idea commonly comprising of utilization and utilization estimations designs of segment appliances over a period interim. Such attempt, in any case, is extremely testing since it is difficult to identify utilization conditions among different appliance at the point when their operation cover or happen in the meantime. Moreover, inferring precise expectation of human movement designs is impacted by the probabilistic connections of machine utilization occasions that have dynamic time interims. To handle the previously mentioned issues, we propose frequent mining and Prediction model to quantify and investigate vitality utilization changes started by inhabitants' conduct. The information from smart meters is recursively mined in the quantum/information cut of 24 hours, and the outcomes are kept up crosswise over progressive mining works out. We likewise use the Bayesian network, a probabilistic graphical model, to anticipate the utilization of numerous machines and household power utilization.

For the assessment of the proposed component, this examination employments the UK Domestic Appliance Level Electricity dataset (UKDale) - time arrangement information of energy utilization gathered from 2012 to 2015 with time determination of six seconds for five houses with 109 machines from Southern England. It must be noticed that by and by stack disaggregation is conveyed by Non-Intrusive Appliance Load Monitoring (NALM) strategy. NALM is a strategy used to disaggregate a home's control use into singular machines and mark them for additionally mining and investigation.

2. OBJECTIVE

Health care services are one of the most challenging aspects that are greatly affected by the vast influx of people to city centers. The underlying correlation between appliance usage inside the smart home and routine activities can be used by health care applications to detect potential health problems. The main goal is to learn occupants' behavioral characteristics as an approach to understand and predict their activities that could indicate health issues. However, the existing system does not consider appliance level usage details, which is critical in identifying human activity variations.

3. RESEARCH METHODOLOGY

Recently, there has been a developing enthusiasm for utilizing smart home innovations for distinguishing human activity designs for wellbeing checking applications. The primary objective is to take in tenants' behavioral attributes as an approach to comprehend and anticipate their exercises that could demonstrate medical problems. In this segment, we survey existing work in the writing, which utilize smart homes information to break down clients' conduct.

Table 1: Related Work

Title	Methodology
Detecting Activities of Daily Living with Smart Meters-2014	The Semi-Markov-Model (SMM) and an impulse based method that also allows the detection of ADLs
"Detecting household activity patterns from smart meter data," in Intelligent Environments (IE)	An indicative activity model for the recognition of four domestic activities using the Dempster-Shafer theory of evidence.
"Smart meter profiling for health applications,"	Classification techniques to detect abnormal behavior of personal energy usage patterns in the home.
"Smart-energy group anomaly based behavioral abnormality detection,"	Hierarchical probabilistic model-based detection to infer about discovered anomalous behavior.
"Appliance usage prediction using a time series based classification approach,"	Time-series multi-label classifier to forecast appliance usage based on decision tree correlations.

4. DATA ANALYSIS

The proposed demonstrate begins by cleaning also, setting up the information and after that applying frequent pattern mining for finding appliance to-appliance affiliations, i.e., figuring out which appliance are working together. At that point, it utilizes group investigation to decide appliance-to-time affiliations. With these two procedures, the framework is ready to extricate the example of appliance utilization which is at that point utilized as input to the Bayesian Network for short-term and lo-term activities forecast. The yield of the framework is used by particular healthcare services applications relying upon the planned utilize. For instance, a social insurance supplier may just inspired by knowing exercises identified with psychological hindrance where following the grouping of every day exercises is essential for reminding the patient when anomalous conduct is recognized.

A model that utilizes smart home big data as a means of learning and discovering human activity patterns for health care applications. Since occupant's habits are mostly identified by everyday routines, discovering these routines makes it possible to recognize anomalous activities that may indicate occupant's difficulties in taking care for themselves. The proposed system addresses the need to analyze temporal

energy consumption patterns at the appliance level, which is directly related to human activities.

The following listed strategies are used to measure and analyze energy usage changes sparked by occupants' behavior to understand and predict household activities that could indicate health issues.

- Incremental frequent pattern mining,
- K-means cluster analysis and
- Bayesian Network activity prediction

Figure 1: Model : Mining Frequent Patterns and Activity Predictions for Healthcare Applications in Smart Homes

5. EXPERIMENTAL RESULTS

For the evaluation of the proposed model, experiments are performed using the dataset UK-Dale along with the synthetic dataset to inspect intermediate and final results. The (UK-Dale) dataset includes time series data of power consumption collected between 2012 and 2015. The dataset contains time series data for five houses with a total of 109 appliances, having a time resolution of 6 seconds, from Southern England published by UK Energy Research Centre Energy Data Centre (UKERC-EDC). This dataset is one of the largest datasets having approximated half a billion records. Energy consumption measurement was conducted at appliance level using plug-in individual appliance monitors (IAMs). The underlying system for the proposed model is developed in Hadoop, and the data is stored in HDFS on Ubuntu 14.04 LTS 64-bit system.

The main objective of the experiments is to detect the appliance usage as an indication of human activity patterns and use the prediction model to forecast the short and long term activities inside the house. For a health care application, this means that our model can be used to feed mechanisms such as active monitoring, alert generation, health profiling etc.

The first step in understanding human activities is by extracting associations of appliance usage. Figure (2) and (3) show the appliance-to-time associations discovered for time of the day and weekday respectively for house 2.

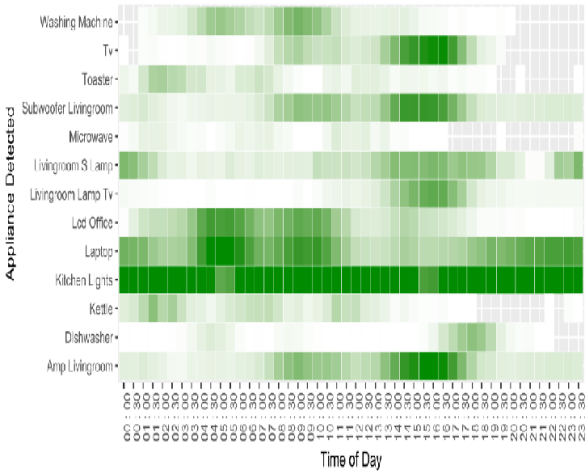

Figure. 2: Appliance-time associations hour of the day

We can see that between 2:30 and 5:00 PM TV, Toaster, Living room Lights are used together in this house with highest concentration during the weekend. Also, the washing machine and Laptop are simultaneously used between 8:30 and 10 am. The washing machine is used almost all weekdays, where the Laptop is not used on the weekends. Considering these facts we can notice the varying effect of time and days on the use of appliances.

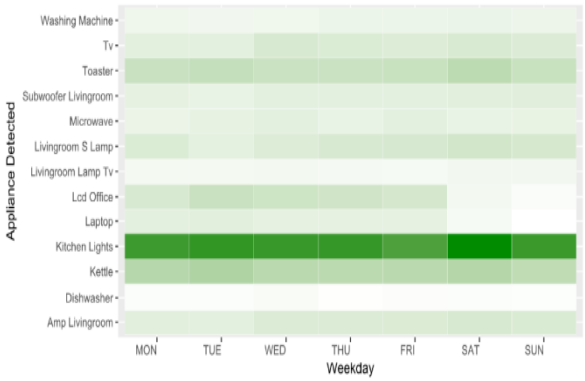

Figure. 3: Appliance-time associations day of the week

One can easily observe from appliance associations that occupants of house1 like to relax while preparing food. This is evident from strong associations of appliances Kitchen Lights, Subwoofer, Amp, and TV. For house 2, occupants like to use the computer or listen to music while washing clothes. Similarly for house number 5, occupants use the computer while cooking or doing the laundry. Such associations are characteristics of human behavioral traits that are performed routinely.

Table (2) shows few examples of possible activities inside

the house based on the probability of appliance associations. These are just samples of human activities that can be discovered by our system and be used to detect anomalies that deviate from normal patterns.

Table 2 : Examples of activity recognition results

No	Time	Appliances detected	Probability	Activity
1	5:00-7:00	Kitchen lights, Toaster, Kettle	20.794	Preparing Breakfast
2	9:00-11:00	Laptop, Lcd Office, Washing machine	60.45	Using Computer while doing laundry
3	13:00-14:00	Microwave, Laptop, TV	24.47	Relaxing
4	17:00-18:00	TV, Washing machine, Living room lights	38.43	Watching TV while doing laundry

The prediction model utilizes appliance-to-appliance and appliance-time associations to predict multiple concurrent operating appliances. We can easily see the strong relationship between appliance usage inside the smart houses and human activity recognition. Learning the appliance-to-appliance and appliance-to-time associations extracted from the frequent pattern mining and cluster analysis are key processes to track patients/people's routines and possibly provide them with health services when needed.

6. ACKNOWLEDGEMENT

The authors are grateful to the Department of Computer Science, Muffakhamjah College of Engineering & Technology for supporting through the project work.

7. CONCLUSION

A model for recognizing human activities patterns from low resolution smart meters data has been presented. The main objective of the experiments is to detect the appliance usage as an indication of human activity patterns and use the prediction model to forecast the short and long term activities inside the house.

Learning the appliance-to-appliance and appliance-to-time associations extracted from the frequent pattern mining and cluster analysis are key processes to track patients/people's routines and possibly provide them with health services when needed.

The model is refined to incorporate distributed learning of big data mining from multiple houses in a near real-time manner. For a health care application, this means that the developed model can be used to feed mechanisms such as active monitoring, alert generation, health profiling etc.

8. REFERENCES

[1] N. United, "World urbanization prospect." United Nation, 2014.[Online].Available:http://dl.acm.org/citation.cfm?id=308574. 308676

[2] M. S. Hossain, "Cloud-supported cyber-physical localization framework for patients monitoring," IEEE Systems Journal, vol. 11,no. 1, pp. 118–127, March 2017.

[3] M. S. Hossain, G. Muhammad, W. Abdul, B. Song, and B. Gupta,"Cloud-assisted secure video transmission and sharing framework for smart cities," Elsevier, Future Generation Computer Systems Journal, April 2017.

[4] J. Liao, L. Stankovic, and V. Stankovic, "Detecting household activity patterns from smart meter data," in Intelligent Environments (IE), 2014 International Conference on, 6 2014, pp. 71–78.

[5] A. Yassine, A. A. N. Shirehjini, and S. Shirmohammadi, "Smart meters big data: Game theoretic model for fair data sharing in deregulated smart grids," IEEE Access, vol. 3, pp. 2743–2754, 2015.