Mining Human Activity Patterns from Smart Home Big Data for Health Care Applications

[1]P.SAKTHIVEL, [2]J.VINOTHINI, [3]P.VIMALA, [4]J.MAHESWARI

[1, 2,3,4] Assistant Professor/CSE
Dhirajlal Gandhi College of Technology, Salem, Tamilnadu.
sakthivelp.cse@dgct.ac.in
vinothini.cse@dgct.ac.in
vimala.cse@dgct.ac.in
mahajm.cse@dgct.ac.in
[1]9944103952

Abstract— Nowadays, there is an ever-increasing migration of people to urban areas. Health care service is one of the most challenging aspects that is greatly affected by the vast influx of people to city centers. Consequently, cities around the world are investing heavily in digital transformation in an effort to provide healthier ecosystems for people. In such a transformation, millions of homes are being equipped with smart devices (e.g., smart meters, sensors, and so on), which generate massive volumes of fine-grained and indexical data that can be analyzed to support smart city services. In this paper, we propose a model that utilizes smart home big data as a means of learning and discovering human activity patterns for health care applications. We propose the use of frequent pattern mining, cluster analysis, and prediction to measure and analyze energy usage changes sparked by occupants' behavior. Since people's habits are mostly identified by everyday routines, discovering these routines allows us to recognize anomalous activities that may indicate people's difficulties in taking care for themselves, such as not preparing food or not using a shower/bath. This paper addresses the need to analyze temporal energy consumption patterns at the appliance level, which is directly related to human activities. The data from smart meters are recursively mined in the quantum/data slice of 24 h, and the results are maintained across successive mining exercises. The results of identifying human activity patterns from appliance usage are presented in detail in this paper along with the accuracy of short- and long-term predictions.

Keywords: Big data, smart cities, smart homes, health care applications, behavioral analytics, frequent pattern, cluster analysis, incremental data-mining, association rules, prediction.

I Introduction

The demand for health care resources will be greatly affected by this vast influx of people to city centers. This unprecedented demographic change places enormous burden on cities to rethink the traditional approaches of providing health services to residents. In responding to the new needs and challenges, cities are currently embracing massive digital transformation in an effort to support sustainable urban communities, and provide healthier environment. In such transformation, millions of homes are being equipped with smart devices (e.g. smart meters, sensors etc.) which generate massive volumes of fine-grained and indexical data that can be analyzed to support health care services. Advancement of big data mining technologies, which provide means of processing huge amount of data for actionable insights, can aid us in understanding how people go about their life.

We propose a human activity pattern mining model based on appliance usage variations in smart homes. The model which utilizes FP-growth for pattern recognition and k-means clustering algorithms is capable of identifying appliance-to-appliance and appliance-to- time associations through incremental mining of energy consumption data. This is not only important to deter- mine activity routines, but also, when utilized by health care application, is capable of detecting sudden changes of human activities that require attention by a health provider.

We apply a Bayesian network for activity prediction based on individual and multiple appliance usage. This is significant for health applications that incorporate reminders for patients to perform certain activities based on historical data. For added accuracy of the system, the prediction model integrates probabilities of appliance- to-appliance and appliance-to-time associations, thus recognizing activities that occur in certain patterns more accurately.

II Related Work:

The main goal is to learn occupants' behavioral characteristics as an approach to understand and predict their activities that could indicate health issues. In this section, we review existing work in the literature, which employ smart homes data to analyze users' behavior. Detecting human activities in smart homes by means of analyzing smart meters data. The paper proposes two approaches to analyze and detect user's routines. One approach uses Semi-Markov-Model (SMM) for data training and detecting individual habits and the other approach introduces impulse based method to detect Activity in Daily Living (ADL) which focuses on temporal analysis of activities that happen simultaneously. Similarly, the work proposes human activity detection for wellness monitoring of elderly people using classification of sensors related to the main activities in the smart home. Smart meters data are also used for activity recognition using Non-intrusive Appliance Load Monitoring (NALM) and Dempster-Shafer (D-S) theory of evidence.

The study collects pre-processed data from homes to determine the electrical appliance usage patterns and then employs machine learning-based algorithm to isolate the major activities inside the home. The issue is that the study has to perform two steps on the data to completely isolate the main activities. Exploiting appliance usage pat- terns and identify them for sudden behavioral change is presented. The aim of the study is to provide around the clock monitoring system to support people's suffering from Alzheimer or Parkinson disease at minimum intrusion level.

The study uses classification techniques to detect abnormal behavior of personal energy usage patterns in the home. Other studies although do not utilize smart meters data; they use Internet of Things

(IoT) infrastructures in smart cities for developing applications that monitor and provide health services for patients.

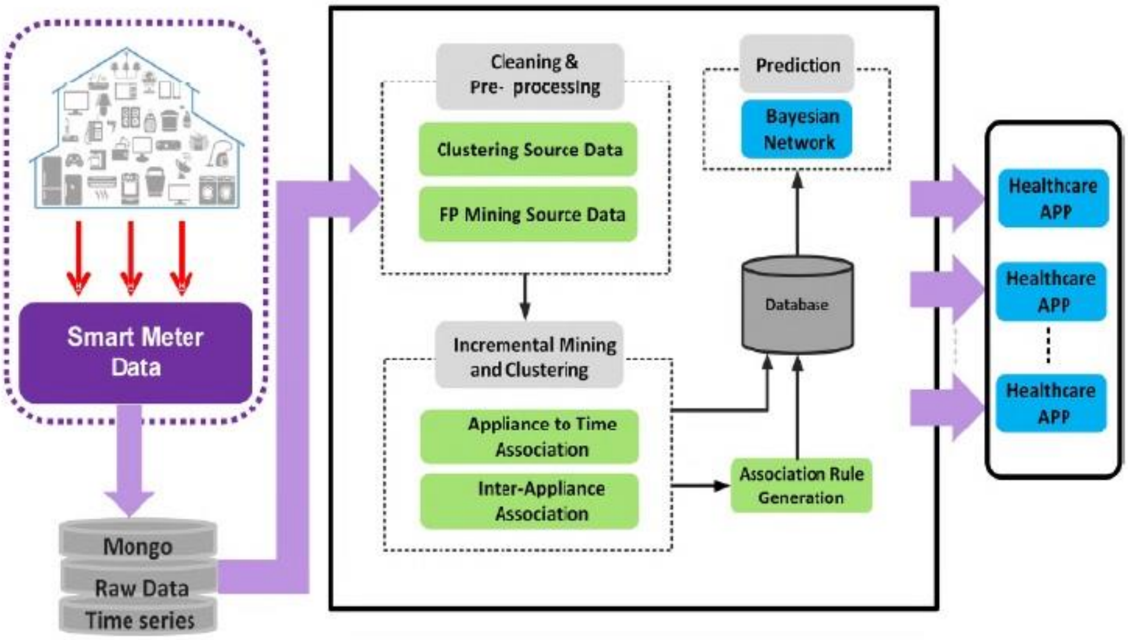

FIGURE 1. Model: Mining frequent patterns and activity predictions for health care applications in smart homes.

This figure and the algorithm is taken from the paper Mining human activity patterns from smart home big data for health care applications proposed by Abdulsalam Yassine et.al. This algorithm outlines the incremental frequent pattern mining process. It requires two kinds of databases like transaction database (DB) and frequent pattern discovered database (FP_DB). The DB stores the source data and FP_DB stores the frequently occurring patterns in the source data. During the incremental frequent pattern mining process it has to be ensured that while frequent patterns are discovered, it should be stored in FP_DB. The steps of the algorithm go like this. Initially for all transaction data in DB, the data has to be processed in the quanta of 24 hours. Then determine database size. Next mine the frequent patterns using the extended FPgrowth approach. Further for all the frequent patterns found in the time slice of 24 hours, search for a frequent pattern in FP_DB. If a frequent pattern is found then update the frequent pattern in FP_DB or else if a new frequent pattern is found then add that frequent pattern to FP_DB. For all the frequent patterns in FP_DB, the database size has to be incremented by the size of the database for the quanta of 24 hours.

A. Detecting Activities of Daily Living with Smart Meters

Detecting activities of daily living with smart meters is a research work in which smart meters are used to provide information to analyze the energy consumption of buildings and to identify the usage of appliances. This helps the older people to stay longer independent in their homes by detecting their activity and their behavior models to ensure their healthy level. This paper can be used to analyze smart meter data to monitor human behavior in single apartments. There are two approaches focused by this paper. They are Semi Markov Model (SMM) and Influence based method. The Semi-Markov-Model (SMM) is used to analyze and detect individual habits to find unique structures representing habits.

If the most possible executed activity (PADL) is evaluated then it can infer the currently executed activity (ADL) of the inhabitant. The impulse based method is used for the detection of ADLs by analyzing all parallel ADLs. Both approaches are based on smart meter events which help to detect which home appliance was switched. Thus, this paper will also give an overview of popular methods to detect the events on electricity consumption data.

B. The Elderly's Independent Living in Smart

Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development

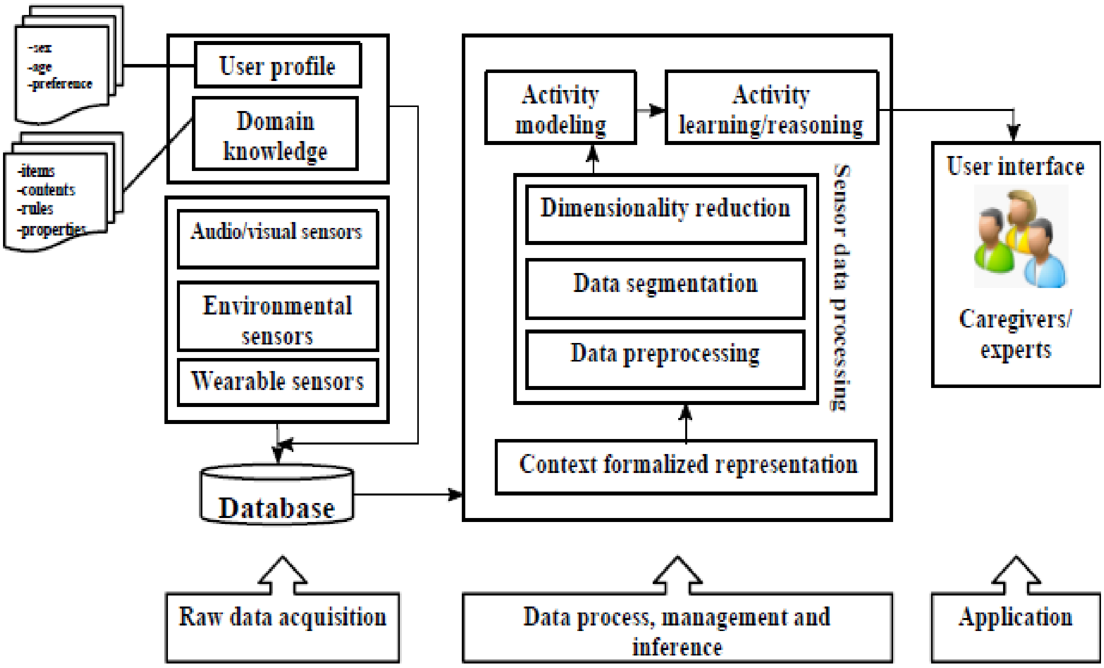

FIGURE 2. Stages in AAL

This is the architecture taken from the paper Elderly's independent living in smart homes: a characterization of activities and sensing infrastructure survey proposed by Q.Ni.A.B.G .Hernando et.al to facilitate services development. Such a living is termed as Ambient Assisted Living (AAL).The activity based AAL consists of three stages: Raw data acquisition, Sensor data processing and learning/reasoning by caregivers. In raw data acquisition stage, the user profile which consists of details of user like age and domain knowledge like items inside smart home are stored in database. There are various sensors like audio/visual, environmental and wearable sensors inside the smart home that collects data and this sensor data are also stored in the database. Next stage is sensor data processing in which the data from the database is taken and transformed into a context formalized representation. Now this data is preprocessed in order to remove noise. The preprocessed data is segmented to partition the data into groups of data having similar properties. The segmented data will undergo dimensionality reduction in which the dimensions of the data are reduced such that it is transformed into a form appropriate for mining. Then an activity modeling occurs in which a model is created based on the human activities inside the smart home. Further these activities are learnt by the caregivers/experts through the user interface of the health care applications in order to detect health problems of humans inside the smart homes.

III Proposed Model:

It starts by cleaning and preparing the data and then applying frequent pattern mining for discovering appliance-to-appliance associations, i.e., determining which appliances are operating together. Then, it uses cluster analysis to determine appliance-to-time associations. With these two processes, the system is able to extract the pattern of appliance usage which is then used as input to the Bayesian network for short-term and long-term activities prediction. The output of the system is utilized by specific health care applications depending on the intended use. For example, a health care provider might only interested in knowing activities related to cognitive impairment where tracking the sequence of daily activities is crucial for reminding the patient when abnormal behavior is detected.

A. DATA PREPARATION

The dataset used in this study is a collection of smart meters data from five houses in the United Kingdom (UK). This dataset includes 400 million raw records at time resolution of 6 seconds. In the first stage of the cleaning process we developed customized procedures to remove noises from the data and prepare it for mining. After cleaning and preparation, the dataset is reduced to 20 million. Additionally, we developed a synthetic dataset for preliminary evaluation of the model, having over 1.2 million records. In tables 1 and 2 we show an example of the resulting ready to mine source data format comprising four appliances from one house. Smart meters time-series raw data, which is a high time-resolution data, is transformed into a 1-minute resolution load data; subsequently translated into a 30 minutes time-resolution source data, i.e. 24 _ 2 D 48 readings per day per appliance, while recording start time and end time for each active appliance.

B. EXTRACTING FREQUENT PATTERNS OF HUMAN ACTIVITIES

As mentioned earlier, the aim is to discover human activity patterns from smart meters data. For example, activities such as ``Watching TV, Cooking, Using Computer, Preparing Food and Cleaning Dishes or Clothes" are usually regular routines. Our aim is to detect the patterns of these activities so that a health care application, that monitors sudden changes in patient's behavior (e.g. patients with cognitive impairment), can send timely alert to health care providers. In pursuing such process, all appliances that are registered active during the 30-minute time interval are included into the source database for frequent pattern data mining.

The energy trace of appliances (TV, Oven and Treadmill) is related to human activities such as leisure/relaxation time, food preparation, and exercising. A simplified example which describes possible relationships between appliance usage and activities is shown in figure 3. Extracting human activity patterns is not only discovering the individual appliance operation, but also the appliance-to-appliance associations; i.e., the patterns of activities that are combined together such as washing clothes while exercising or watching TV.

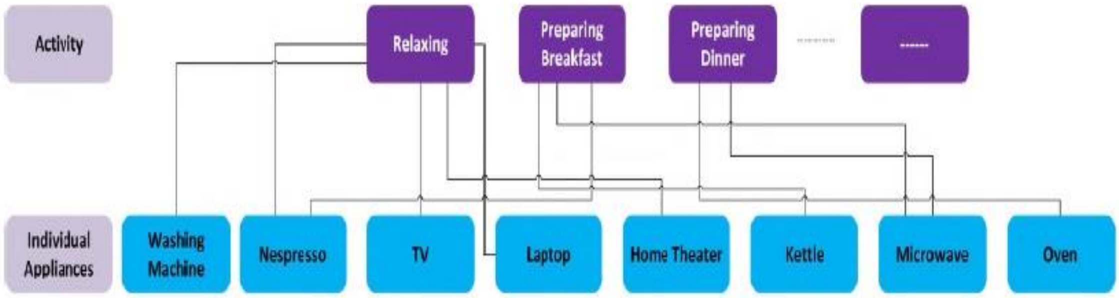

FIGURE 3. Example of possible relationships between appliance usage and daily activities inside typical smart home

V Conclusion:

In this paper, we presented a performed for monitoring the human activities inside a smart home which can be utilized by health care applications to detect health problems. It was found that the former research works does not consider appliance level patterns which are a critical factor to determine human activity variations. This paper proposes a model which is used for recognizing human activities patterns from smart meters data. The human habits and behavior follow a pattern that could be used in health applications to track the health problems of individuals living alone or those with self-limiting conditions. These human activities can be inferred from appliance-to- appliance and appliance-to-time associations. An incremental frequent mining and prediction model is proposed based on Bayesian network. In the proposed work, 24-hour period was found to be optimal for data mining, but the model can operate on any quantum of time. The applicability of the proposed model was to correctly detect multiple appliance usage and make short and long term prediction. The output of the system is utilized by specific health care applications depending on the intended use. For example, a health care provider might only be interested in knowing activities related to cognitive impairment where tracking the sequence of daily activities is crucial for reminding the patient when abnormal behavior is detected.

REFERENCES

[1] United Nation. (2014). World Urbanization Prospect. [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 Syst. J., vol. 11, no. 1, pp. 118_127, Mar. 2017.

[3] M. S. Hossain, G. Muhammad, W. Abdul, B. Song, and B. B. Gupta, ``Cloud-assisted secure video transmission and sharing framework for smart cities,'' Future Generat. Comput. Syst. J., Elsevier 2017, to be published. [Online]. Available: https://doi.org/10.1016/j.future.2017.03.029

[4] J. Liao, L. Stankovic, and V. Stankovic, ``Detecting household activity patterns from smart meter data,'' in Proc. Int. Conf. Intell. Environ. (IE), vol. 6. Jul. 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.

[6] A. Yassine and S. Shirmohammadi, ``Measuring users' privacy payoff using intelligent agents,'' in Proc. IEEE Int. Conf. Comput. Intell. Meas. Syst. Appl., May 2009, pp. 169_174.

[7] A. Yassine and S. Shirmohammadi, ``A business privacy model for virtual communities,'' Int. J. Web Based Commun., vol. 5, no. 2, Mar. 2009.

[8] Y. C. Chen, H. C. Hung, B. Y. Chiang, S. Y. Peng, and P. J. Chen, ``Incrementally mining usage correlations among appliances in smart homes,'' in Proc. 18th Int. Conf. Netw.-Based Inf. Syst. (NBiS), 2015, pp. 273_279.

[9] K. Jack and K.William, ``The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from _ve UK homes,'' Sci. Data, vol. 2, p. 150007, Sep. 2015.

[10] J. Clement, J. Ploennigs, and K. Kabitzsch, ``Detecting activities of daily living with smart meters,'' in Advance Technology and Societal Change. Heidelberg, Germany: Springer, 2014, pp. 143_160. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-37988-8_10