IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

A Survey on Large Scale Data Analysis on Human Activity Patterns for health prediction

P. Geethanjali

Department of Computer Science and Engineering,
Gokaraju Rangaraju Institute of Engineering and Technology, Hyderab

DOI: https://doi.org/10.26438/ijcse/v7i5.12281231 | Available online at: www.ijcseonline.org

Accepted: 25/May/2019, Published: 31/May/2019

Abstract - In this exploration work, big data gathered from smart devices have been utilized to recover the human activity patterns to enhance smart home occupant's health status, as there is a great deal of financial investment in the advanced transformation as a push to give healthier biological communities to individuals. 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.In this transformation a more of smart-devices are prepared around and gives an arranged data that can be utilized to investigate the health data. In this examination, the work mostly centers on breaking down the big data separated from human activities for frequent pattern mining, cluster analysis, prediction to quantify and investigate the energy consumption changes likewise by inhabitants. This paper speaks to the need of breaking down energy-consumption pattern dependent on the machine level, which is totally identified with person behavior.

Keywords: Smart Devices, Human Activity Patterns, Smart home, Cluster Analysis, Bayesian network.

I. INTRODUCTION

The intrigue for the health care assets is by and large generally overpowered by the digital transformation. As indicated by study, continuously 2050 digital transformation assumes a vital job. By progression of these machines, large portion of data created from smart devices [1]. For instance, evaluating the progressions of appliance use can be utilized for aberrant assurance of the individual's prosperity dependent on the historical data. Ordinary routines and activities mirrors their regular habits, on watching their regular habits individuals troubles in dealing with themselves, likely not washing his clothes, not utilizing woven and assist us with recognizing any peculiar activities which may be a sign of sick Health[4].

A technique for the utilization of energy data, gathered from smart devices introduced at homes, and gets data dependent on the routine-activities of occupants. Scholarly investigation of that data can discover the adjustments in
conduct or in health of the inhabitants [3]. Power consumption and the time of use are firmly identified with the resident's activities performed at family unit. For example, if the "Broiler" is ON, the task of this appliance is in all likelihood managing activity "Getting ready Food". Besides, individuals frequently perform more than one activity in the meantime [4]. We break down, energy consumption patterns at the appliance level to foresee their tasks by distinguishing numerous appliance usages. Be that as it may, it is exceptionally testing since it isn't so natural to recognize utilization conditions among different appliances when their activity happens in the meantime [6].
To deal with the recently referenced issues, The information from savvy meters is recursively found in the quantum/information segment of 24 hours, and the results are shielded transversely over coming about mining exercises.

II. SURVEY ON ANALYSIS OF HUMAN ACTIVITY PATTERNS

In 2016 M. S. Hossain proposed a work through paper for appearing patient's state acknowledgment framework for healthcare utilizing discourse and outward appearance. This paper discloses a model to address a general framework on health care. It for the most part manages the idea of

recognizing a patient state for giving great acknowledgment exactness to give minimal effort modeling. This paper basically relies upon two kinds of information sources extensively sound and video which are caught in a multi-sensory condition which demonstrated a normal discovery proficiency more than 98%.

In 2016 M. UlAlam proposed their work on smart-energy aggregate contrast dependent on behavioral peculiarity recognition. Here a data systematic access that arranges energy uses abnormality according to the behavioral deformity of the occupant. Research work for the most part depends on distinguishing everyday appliances utilization from smart meter and smart attachment data tracks regular activity at days and evenings, then learning the one of a kind time fragment gathering of every appliance energy consumption.

In 2014 J. Merciful, J. Ploennigs, and K. Kabitzsch displayed their work in paper for recognizing day by day living activities with smart meters. This paper clarifies the techniques that are utilized to investigate smart meter data to monitor human behavior. A SMM is utilized to break down and discover select structures which characterize habits of unit. The second methodology depends on a form that permits the revelation of ADL‟s and spotlights on temporal pursuit of ADL‟s.

In 2015 K. Gajowniczek and T. Za_bkowski proposed a work in data mining strategies for distinguishing family unit attributes dependent on smart meter data. The objective of the paper is to regulate the structure of family unit appliances use patterns, henceforth giving more acumen in smart metering systems by in perspective of the utilization and the spell of their use. The work conveys the arrangements reasonable for smart metering systems that may add to sophisticated energy awareness; support exact utilization foreseeing and give the contribution to demand reaction systems in family units with periodical energy sparing proposals for clients. This paper conveyed a few outcomes which demonstrate that characterizing family unit features from smart devices.

In 2016 S. Singh, A. Yassine Here Incremental mining of daily power usages from smart meters big data. This paper clarifies; the energy use performance replicates appliance associations and their utilization. The consumption of data from a smart meter is a steady procedure. After a specific period, illustrates the power consumption of a home that can be estimated by the data comprising of appliance use tuples for 24hours in a dynamic way.

In 2011 J. Han, J. Pei work in Data mining: Concepts and strategies, in Cluster Analysis: normal models, third ed. Bayesian network delivers a graphical network of basic relations from which learning can be accomplished. These prepared Bayesian conviction networks are utilized for the classification. These Classifications should be possible dependent on frequent patterns. These frequent patterns mirror the associations among attribute-esteem sets. In, Associative classification the classification depends on association rules generated from frequent patterns though, semi-administered classification is helpful for enormous elements of unsupervised data.

III. EXISTING CASE STUDIES

In an implementation, the family unit data is gathered from the prepared smart devices to gauge the activity patterns of the living creatures. In the wake of gathering the meter readings, the data is siphoned for pattern mining utilizing activity pattern algorithms.

A succession of steps incorporates doing the work for recognizing the human activity patterns. This work recognizes a model that investigations the human activity patterns of smart homes residents for the health prediction. To execute the model, we gather the smart home data and apply pattern mining algorithms and apply clustering algorithms.

First it starts by applying regular example mining to discover machine to-apparatus relations that understands which machines are cooperating. At that point, this model uses group investigation by seeing k means and DBscan bunching calculations for discovering machine to-time affiliations. With the assistance of over two procedures, the model is sufficiently proficient to derive the pattern of appliance utilization which will be utilized as contribution to the Bayesian network for activities prediction. The yield created by the framework is helpful for explicit health care applications relying upon the anticipated use.

Case Study Implementation Steps

1. Smart homes are outfitted with smart devices.
2. When the family units utilize the smart devices, huge volumes of data are generated.
3. Smart devices utilization resembles watching the routine activities of the family units to perceive peculiar activities.
4. The smart meter generated huge volumes of data are gathered.
5. The data is clustered dependent on the clustering algorithms.
6. After the cluster analysis the source data is refined into regulated learning classification.
7. Then the FP development mining is connected to mine the source data.
8. Then acquired data is gathered and set for the prediction to foresee the human activity patterns dependent on the appliance use.

9. Based on the Bayesian networks we manufacture an activity prediction model which in wording causes us to anticipate the health states of the family units.

Identifying Frequent Patterns in source data

As referenced before, the point is to find human activity patterns from smart meters data. For instance, activities, for example, Watching TV, Cooking are normally regular routines. Our point is to identify the patterns of these activities with the goal that a health care application monitors sudden changes in patient's behavior send timely health care suppliers.

For example, recreation/relaxation time, food preparation, and working out. A streamlined precedent which portrays conceivable relationships between appliance use and activities. Procuring human activity patterns isn't just watching the individual appliance operation, yet additionally the appliance-to-appliance associations that is the patterns of activities that are joined. The fundamental idea of the model depends on which propose pattern development or FP-development approach and Apriori Algorithm for frequent pattern mining.

IV. CLUSTER ANALYSIS

Perceiving appliance-to-time associations has a major job in health applications that keeps a track on inhabitant's activity patterns constantly. In this area, to recognize about the appliance use time clustering analysis method is utilized. Appliance-to-time associations are hidden information in the smart devices time series data which incorporate adequately close time-stamps, when applicable appliance has been recorded as dynamic or operational. The extent of the cluster that depicts such associations is characterized as include of individuals the cluster just as its relative quality. Clustering analysis is the way toward creating classes or gatherings/fragments or parcels where individuals must have closeness with each other, however ought to be different from the individuals from the other clusters. The unmistakable preferred standpoint of the clustering is the non-administered nature of the procedure.

K-Means Clustering Algorithm

A partitioning technique connected to overview data and considers observations of the data as articles dependent on locations and separation between different information data points. Partitioning the articles into usually constrained clusters (K) is finished by it in such a way, to the point that objects inside each cluster endure as close-by as likely to one another.

Each cluster is considered by its Centroid i.e., its middle point. The partings utilized in clustering in the majority of the stages don‟t truly imply the spatial separations. In wide-running, the main goals for this issue of end worldwide least is complete decision of beginning stages. A centroid is where coordinates are procured by means of registering every coordinate of the points, models allocated to the clusters.

DBscan Clustering Algorithm

The principle impression of the DBSCAN algorithm each point of a cluster and neighborhood of a given span must contain a base number of point's i.e the thickness in the neighborhood must stifle some predefined edge. This methodology needs input parameters:

K, the proximate neighbor list size.

1. Eps, the radius that determine the neighborhood region of a point (Eps-neighborhood).
2. MinPts, the minimum number of points that must exist in the Eps-neighborhood.

V. PREVIOUS RESULTS

In this area, K-Means clustering algorithm is contrasted and DBscan clustering. What's more, the use of Bayesian networks for human activity pattern acknowledgment likewise presented.

Table 1 shows the comparison between K-Means and DBscan clustering algorithms.

Parameters	K-Means	DBscan
Approach	Partitional Based	Density based
Characterization	Centroid based	Dense Region based
Limitations	When clusters are of different Size, Densities, Non-globular shapes. When the data contains outliers.	Do not work efficiently when there are more number of clusters with different densities.
Advantage	More quicker compared to DBscan.	No cluster size is is demanded to form clusters.

Smart home data can be utilized with either K-Means clustering algorithm or DBscan clustering algorithm. The choice of algorithm is generally relying upon dataset. On the off chance that number of clusters is predefined and if the dataset is adaptable K-Means can be connected. In the event that no prior information about number of clusters, DBscan is doable.

Here we utilized MOA tool to plot both K-Means and DBscansss algorithms for a dataset. The consequences of both the clustering algorithms can be looked at and dependent on our necessity we can utilize the dataset as a contribution for Bayesian networks.

VI. CONCLUSION

Here Occupants' propensities and conduct seek after an example that could be used in wellbeing applications to follow the thriving of individuals living alone or those with self-confining conditions. By far most of these exercises can be picked up from machine to-apparatus and machine to-time affiliations. Furthermore, moreover can presented unfaltering regular mining and expectation display subject to Bayesian system. In our investigation work, through examinations, we found that 24-hour length was perfect for information mining, yet we manufactured the model to work on any quantum of time.

REFERENCES

[1] Abdulsalam Yassine, Shailendra Singh, and Atif Alamri," Mining Human Activity Patterns from smart home big data for health care applications", IEEE Access, Vol. 5, 2017.

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

[3] K. William and K. Jack, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from UK homes," Sci. Data, Sep. 2015.

[4] M. S. Hossain, "A patient's state recognition system for health care using speech and facial expression," J. Med. Syst., vol. 40, no. 12, Dec. 2016.

[5] M. UlAlam, N. Roy, M. Petruska, and A. Zemp, "Smart-energy group anomaly based behavioral abnormality detection," in Proc. IEEE Wireless Health (WH), Oct. 2016.

[6] K. Kabitzsch, J. Clement and J. Ploennigs "Detecting activities of daily living with smart maters", in Advance technology and Societal Change. Heidelberg, Germany: Springer, 2014.

7] K. Basu, V. Debusschere, and S. Bacha, "Appliance usage prediction using a time series based classification approach," in Proc. IEEE 38th Annu. Conf. Ind. Electron. Soc. (IECON), May 2012.

[8] L. Hawarah, S. Ploix, N. Arghira, H. Joumaa, and K. Basu, "A prediction system for home appliance usage", Energy Buildings, vol.67, Sep. 2013.

[9] V.K. Prasanna, C. Chelmis, and J. Kolte, "Big data analytics for demand response: Clustering over space and time", in Proc. IEEE Int. Conf. Big Data (Big Data), Apr. 2015.

[10] T. Za_bkowski and K. Gajowniczek, "Data mining techniques for detecting House hold characteristics based on smart meter data," Energies, vol. 8, no. 7, 2015.

[11] Y. Yin, R. Rao, J. Pei, and J. Han, "Mining frequent patterns without Candidate generation: A frequent-pattern tree approach", Data Mining Knowl. Discovery, vol. 8, no. 1, 2004.

[12] J. Han, J. Pei, and M. Kamber, "Data mining: Concepts and techniques,"in Cluster Analysis: Basic Concepts and Methods, 3rd ed. San Mateo, CA, USA: Morgan Kaufmann, 2011.

[13] D. Heckerman, "Bayesian networks for data mining," Data Mining Knowl. Discovery, vol. 1, no. 1, 1997.

[14] S. Shrimohammadi, A. Yassine and S. Singh, "Incremental mining of frequent power consumption patterns from smart maters big data", in Proc. IEEE Electrical Power Energy Conf. (EPEC), Oct. 2016..

[15] J. Han, J. Pei, and M. Kamber, "Data mining: Concepts and techniques," in Classification: Advanced Methods, San Francisco, CA, USA: Morgan Kaufmann, 2011.

[16] Mrs. Bhawana Mathur, Dr. Manju Kaushik: "Comparative study of k-means and Hierarchical Clustering Techniques", International journal of Software & Hardware Research in Engg, 2014.

[17] Dr. M Nagalakshmi, Dr. I Surya Prabha, K Anil, Big Data Map Reducing Technique Based Apriori in Distributed Mining. International Journal of Advanced Research in Engineering and Technology, 8(5), 2017, pp 19 – 28.

[18] Parag C. Shukla and Dr. Kishor Atkotiya, Big Data Analytics: What It Is and What It Isn"t, Characteristics, Classification, Challenges and Importance. International Journal of Computer Engineering & Technology, 8(6), 2017, pp. 60–66.

[19] K. Prema and Dr. A.V. Sriharsha, Differential Privacy in Big Data Analytics for Haptic Applications. International Journal of Computer Engineering & Technology, 8(3), 2017, pp. 11–19.

[20] Naga Raju Hari Manikyam and Dr. S. Mohan Kumar, Methods and Techniques To Deal with Big Data Analytics and Challenges In Cloud Computing Environment. International Journal of Civil Engineering and Technology, 8(4), 2017, pp. 669-678.

