VOLUME V
ISBN No.: 978-81-953602-8-4
Computational Science

Estd. 1957

NALLAMUTHU GOUNDER MAHALINGAM COLLEGE

An Autonomous Institution, Affiliated to Bharathiar University, An ISO 9001:2015 Certified Institution,

Pollachi-642001

SUPPORTED BY

Sakthi™ as pure as mother's love

ABT Industries Ltd.,
Dairy Division

Riyasaa Labs

PROCEEDING

One day International Conference

EMERGING TRENDS IN SCIENCE AND TECHNOLOGY (ETIST-2021)

27th October 2021

Jointly Organized by

Department of Biological Science, Physical Science and Computational Science

NALLAMUTHU GOUNDER MAHALINGAM COLLEGE

An Autonomous Institution, Affiliated to Bharathiar University

An ISO 9001:2015 Certified Institution, Pollachi-642001.

Estd. 1957

Proceeding of the

One day International Conference on

EMERGING TRENDS IN SCIENCE AND TECHNOLOGY (ETIST-2021)

27th October 2021

Jointly Organized by

Department of Biological Science, Physical Science and Computational Science

ISBN No: 978-81-953602-8-4

Nallamuthu Gounder Mahalingam College

An Autonomous Institution, Affiliated to Bharathiar University

An ISO 9001:2015 Certified Institution, 90 Palghat Road, Pollachi-642001.

www.ngmc.org

International Conference on Emerging Trends in Science and Technology (ETIST 2021)
Jointly Organized by Department of Biological Science, Physical Science and Computational Science
Nallamuthu Gounder Mahalingam College, Affiliated to Bharathiar University, Tamilnadu, India.
Published by NGMC - November 2021

Analysis of Energy Efficient Power Scheduling in Smart Home Automation System Using Canopy and K-Means Data Mining Clustering Algorithms

Sharmila S[1] - Dr. Antony Selvadoss Thanamani[2] - Finny Belwin A[3] - Linda Sherin A[4] - Kanagaraj A[5] - Dr P.T. Anitha[6]

ABSTRACT: Nowadays, increasing energy efficiency is used to face the great world challenges, such as energy security, air pollution, climate change and economic crises, among others. Energy efficiency alternatives have the power to optimize energy consumption and reduce greenhouse gas emissions, thus contributing positively to the preservation of natural ecosystems and human health. Internet of Things (IoT) is a technology which enables computing devices, physical and virtual devices to be connected to the internet so that users can control and monitor devices. The IoT offers huge potential for development of various applications namely: e-governance, environmental monitoring, military applications, infrastructure management, industrial applications, energy management, healthcare monitoring, home automation and transport systems. In this paper, power consumption on home automation system is compared by implementing algorithms such as canopy algorithm, k-means clustering algorithm and manual system. IoT plays a major role on home automation. The development of IoT based technology for home automation is gaining lot of attention and many researchers are working in that direction. The home automation will help every individual to monitor and control home appliances remotely through handheld devices and computers using internet.

Keywords: Internet of things, data mining, k-means clustering algorithm, canopy algorithm, power consumption, home automation

Sharmila S[1], Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Coimbatore, Tamilnadu, India, Email - mcasharmi2007@gmail.com
Antony Selvadoss Thanamani[2], Associate Professor & Head, Department of Computer Science(Aided), Nallamuthu Gounder Mahalingam College, Pollachi, Coimbatore, Tamilnadu, India, Email - selvdoss@gmail.com
Finny Belwin A[3], Assistant Professor, Department of Computer Science, Angappa College of Arts and Science, Coimbatore, Tamilnadu, Email - belwin35@gmail.com
Linda Sherin A[4], Assistant Professor, Department of Computer Science, AMJAIN College, Chennai, Tamilnadu, Email - linz15sherin@gmail.com
Kanagaraj A[5], Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Coimbatore, Tamilnadu, India, Email - a.kanagaraj@gmail.com
Dr P.T. Anitha[6], Associate Professor - CS and IT, College of Engineering and Technology, Wollega University, Nekemte, Ethiopia, Email: anithapt74@gmail.com

1. INTRODUCTION

In recent years, the Internet has impacted people's daily lives through a new paradigm called IoT, which is present in smart homes, retail, education, government services, smart grids, agriculture, communication and business, among others. The IoT is the combination of various technologies belonging to application domains that interconnect objects or things through the Internet; by doing so, IoT-based things or objects acquire detection, monitoring, and remote-control capabilities. The IoT can collect, distribute, and analyze data to convert it into knowledge and information. Therefore, it is important to further research and develop energy optimization mechanisms across different IoT application domains. Moreover, appropriate energy saving and collection proposals and programming algorithms should continue to be sought, since renewable energy sources are becoming more important every day. Additionally, researchers estimate that the adoption of the IoT is based on the success of these energy optimization proposals.

Smart home is the use of computer technology, control technology, image display technology and communications technology will be connected through the network of various facilities together to meet the automation requirements of the entire system to provide more convenient control and management. The traditional smart home implementation generally controls and communicates building facilities through the wiring of wired lines, it is difficult to get rid of the restraints of various cables, the installation cost is high, and the scalability of the system is also poor. The smart home system based on wireless sensor network technology can not only get rid of the shackles of cables, reduce the installation cost, but also greatly improve the scalability of the system. There are some main features for smart home as follows:

(1) The smart home realizes the real-time meter reading and security service of water meter, electric energy meter and gas meter, which provide more convenient conditions for the high-quality service.
(2) Smart home can enhance the comfort, safety, convenience and interactivity of home life, and optimize people's life style.
(3) Smart home can monitor and interact with the home through telephone, mobile phone and remote network, discover the abnormal and timely processing.
(4) The smart home can realize the interaction between the user and the power grid enterprise, obtain the information of electricity consumption and electricity price, set the electricity consumption plan and so on, guide the scientific and rational electricity use and advocate the family's consciousness of energy saving and environmental protection.
(5) Support "triple networks" business and the perfect intelligent service.
(6) Smart home can support remote payment.

The paper is structured as follows, in section 2 we will provide literature review of various articles related to the home automation, section 3 discussed about data mining, section 4 discussed about clustering algorithms such as canopy and k-means clustering algorithm, section 5 describes about smart home application, section 6 discusses result and analysis followed by conclusion.

1. LITERATURE REVIEW

Nowadays, increasing energy efficiency is used to face the great world challenges, such as energy security, air pollution, climate change and economic crises, among others. Energy efficiency alternatives have the power to optimize energy consumption and reduce greenhouse gas emissions, thus contributing positively to the preservation of natural ecosystems and human health. Additionally, energy efficiency alternatives help mitigate the economic effects in the workplace. Energy efficiency has aroused great interest in research worldwide, because energy consumption has increased in recent years, especially in the residential sector. For this reason, organizations and governments worldwide are proposing actions for energy conservation with the purpose of reducing energy-related inconveniences. The residential sector is attributed a high energy consumption; however, home automation systems (HAS), combined with IoT, are alternatives that promise to contribute to greater energy efficiency.

Home automation systems combined with IoT, big data technologies, and machine learning are alternatives that promise to contribute to greater energy efficiency. In [1] Isaac Machorro-Cano, Giner Alor-Hernández et al., presents HEMS-IoT, a big data and machine learning-based smart home energy management system for home comfort, safety, and energy saving. They used the J48 machine learning algorithm and Weka API to learn user behaviors and energy consumption patterns and classify houses with respect to energy consumption. Likewise, they relied on RuleML and Apache Mahout to generate energy-saving recommendations based on user preferences to preserve smart home comfort and safety. To validate the system, they presents a case study where they monitor a smart home to ensure comfort and safety and reduce energy consumption.

The smart home service is a key part of the smart grid consumption. It is a real-time interactive response between the power grid and users, and enhances the comprehensive service capability of the power grid, also realizes the intelligent and interactive use of electricity, further improves the operation mode of the power grid and the users' Use patterns to improve end-user energy efficiency. The smart home is a residential-based platform that uses IoT, computer technology, control technology, image display technology and communication technology to connect various facilities through the network to meet the automation requirements of the entire system and provide more convenient control and management. In [2] Min Li, Wenbin Gu et al., analyzes the characteristics of smart home, gives the smart home composition and the application of key equipment; and smart home key technologies to illustrate the design of smart home electricity service system and related communication systems.

Power management has been of particular interest as it allows for a greener future and is added advantage to users as a cost-cutting measure. In [3] Dhiren Tejani, Ali Mohammed A. H. Al-Kuwari et al., analyses the power consumption in a standard home taking into account commonly used appliances and other devices and shows the management capabilities of a home automation system as well as an estimation of the savings in terms of power consumed and cost. The aim of the study is to measure the energy conservation across 4 homes using a Z-Wave home automation system. It is found that there is an 18.70% decrease in energy consumption when the home automation system acts to manage the power consumption of the devices in the home.

In [4] Adela Bara, Bogdan Tudorica et al., analysis Big Data storage and processing solutions applicable in the smart grid context in case of large volume of data coming from sensors (smart meters, IoT appliances). A flexible architecture for data management is proposed that consists in three layers: relational database (RB) tier, big data (DG) tier and data warehouse (DW) tier. A proof of concept implementation is also provided, using Elasticsearch and Kibana.

In [5] Shen Bin, Liu Yuan et al., proposed four data mining models for the Internet of Things, which are multi-layer data mining model, distributed data mining model, Grid based data mining model and data mining model from multi-technology integration perspective. Among them, multi-layer model includes four layers: 1) data collection layer, 2) data management layer, 3) event processing layer, and 4) data mining service layer. Distributed data mining model can solve problems from depositing data at different sites. Grid based data mining model allows Grid framework to realize the functions of data mining. Data mining model from multi-technology integration perspective describes the corresponding framework for the future Internet. Several key issues in data mining of IoT are also discussed.

Internet of Things (IoT) has been growing rapidly due to recent advancements in communications and sensor technologies. Interfacing an every object together through internet looks very difficult, but within a frame of time Internet of Things will drastically change our life. The enormous data captured by the Internet of Things (IoT) are considered of high business as well as social values and extracting hidden information from raw data, various data mining algorithm can be applied to IoT data. In [6] Akshat Savaliya, Aakash Bhatia et al., made a survey on systematic review of various data mining models as well as its application in Internet of Thing (IoT) field along with its merits and demerits.

2. DATA MINING

Data mining process refers to the process of semi-automatically analyzing large databases for pattern mining which are innovative, legitimate, useful and understandable which is also known as Knowledge Discovery in Databases (KDD). Data mining or KDD process includes problem formulation, data collection, data cleaning i.e. preprocessing, transformation, choosing mining task and result evaluation. Knowledge discovery is an iterative process. Data mining overlaps with other fields like statistics, machine learning, artificial intelligence, databases but mainly it focuses on automation of handling large heterogeneous data, algorithm and scalability of number of features and instances.

There are various issues involved in data mining in Internet of Things:

a) Efficiency in data gathering - Energy efficiency, scalability and fault tolerance should be taken into consideration when data is to be collected from distributed sensor networks.
b) Data abstraction and aggregation - Managing massive data generated from IoT is a challenging task. Efficient mechanism should be adopted for data deduplication.

c) Distributed data processing and mining - Due to nodes' constraints, paradigm shift is needed for prior level preprocessing of the data at each distributed nodes and an aggregated information is to be sent to sink node in order to optimize energy usage instead of sending all distributed data to server for processing.

d) Data mining towards the next age of Internet - In an upcoming generations of Internet, latest trends and technologies like ubiquitous computing, semantic web, IPv6 technologies are going to be integrated with IoT. This will give rise to challenges for Data Mining due to heterogeneous unstructured data.

Data mining goes through several steps. They are divided into a data conditioning or data preprocessing phase and a predictive analysis phase. In the data conditioning phase, data is collected and preprocessed. Not all data is useful for a specific DM task so selecting the observation points is an important preprocessing step. Other preprocessing steps include data transformation to have consistent data formats, data deduplication and outlier removal. In the predictive analysis phase, suitable DM methods have to be selected and trained. Depending on the problem, the data has to be correlated since data from only one data source might not be enough to make meaningful predictions. Often data is also visualized and reports are generated. DM is highly iterative and some steps might be run through many times.

3. CLUSTERING ALGORITHMS

3.1 K-MEANS

K-means is a standard partitioning clustering method based on K centroids of a random initial partition which is iteratively improved. It is one of the simplest unsupervised learning methods among all partitioning based clustering methods. It aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Clusters generated by k-means are a non-hierarchical. It is well suited to generating Global Clusters. The K-Means method is numerical, unsupervised, non-deterministic and iterative.

K-means is a simple algorithm that has been adapted to many problem domains. As it is going to be seen, it is a good candidate for extension to work with fuzzy feature vectors. As such, it has been successfully used in various areas ranging from Market Segmentation, Computer Vision, Geo Statistics, and Agriculture to Astronomy. It is often used as a pre-processing step for other algorithms, for example to find a starting configuration. Due to its ubiquity it is often called the k-means algorithm; it is also referred to as Lloyd's algorithm, particularly in the computer science community.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

The algorithm follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a prior. The algorithm is composed of the following steps.

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

Initially k-means algorithm classifies a given set of n data objects into k clusters, where k is the number of desired clusters and it is required in advance. A centroid is defined for each cluster. All the data objects are placed in a cluster having centroid nearest (or most similar) to that data object. After processing all data objects, k-means, or centroids, are recalculated, and the entire process is repeated. All data objects are bound to the clusters based on the new centroids. In each iteration centroids change their location step by step. In other words, centroids move in each iteration. This process is continued until no any centroid move. As a result, k clusters are found representing a set of n data objects.

3.2 CANOPY

Canopy clustering algorithm is a pre-clustering algorithm. In general, it is used as a preprocessing step for traditional clustering algorithm. Clustering large datasets with traditional clustering algorithm may take more time to provide results. In contrast if canopy algorithm is used as a preprocessing step moreover half of the process is finished by canopy algorithm. So it took less time to provide result. The canopy based clustering algorithm works in two steps; initially the datasets are partitioned into overlapping subsets called canopies, and then clustering process has been performed on subsets.

Canopy algorithm starts with a list of values in any order and with use of two thresholds T_1 & T_2, where $T_1 > T_2$ for fast distance measurement process. In the first step, canopy algorithm randomly selects one value p among dataset S for cluster center c. Then it will remove p from dataset S and continue the process with the remaining dataset elements. Then distance measure will be calculated between canopy center c and all other elements in dataset S. For each element p_i, if calculated distance is less than threshold value T_1, then element p_i is added to the canopy C_j. Since if calculated distance value for element p_i is less than threshold value T_2 then element p_i is removed from dataset S. Finally, canopy C_j is added to the canopy list C. Same like this, up to the dataset S becomes empty, canopy algorithm repeats this process continuously.

***S**= set of values;*
*While dataset **S** is not empty*
    *select cluster_point **p** from dataset **S** to initialize canopy_center **c***
    *remove cluster_point **p** from **S***
    *Loop through remaining elements in dataset **S***
        *if distance between element $\boldsymbol{p_i}$ and canopy_center **c** < Threshold $\boldsymbol{T_1}$*

then add element p_i *to canopy* C_j

if distance between element p_i *and canopy_center* c < *Threshold* T_2

then remove element from dataset S

End

add canopy C_j *to the list of canopies* C

End

4. SMART HOME AUTOMTATION SYSTEM

Home automation systems have been steadily gaining popularity especially in homes and office spaces. The systems are installed with the intention of providing the user with easier access and more control over the devices in the home. The user will be able to monitor and control devices at home locally as well as remotely via the internet provided the users has a device with a compatible web browser.

4.1 ENERGY EFFICIENT POWER SCHEDULING

Energy efficiency has aroused great interest in research worldwide, because energy consumption has increased in recent years, especially in the residential sector. The advances in energy conversion, along with new forms of communication, and information technologies have paved the way for what is now known as smart homes. The Internet of Things (IoT) is the convergence of various heterogeneous technologies from different application domains that are used to interconnect things through the Internet, thus allowing for the detection, monitoring, and remote control of multiple devices.

5. RESULT AND ANALYSIS

Graph shown in figure 1 depicts various levels of energy consumption of home appliances. The result of power consumption comparison is made between Manual, K-Means and Canopy clustering algorithms.

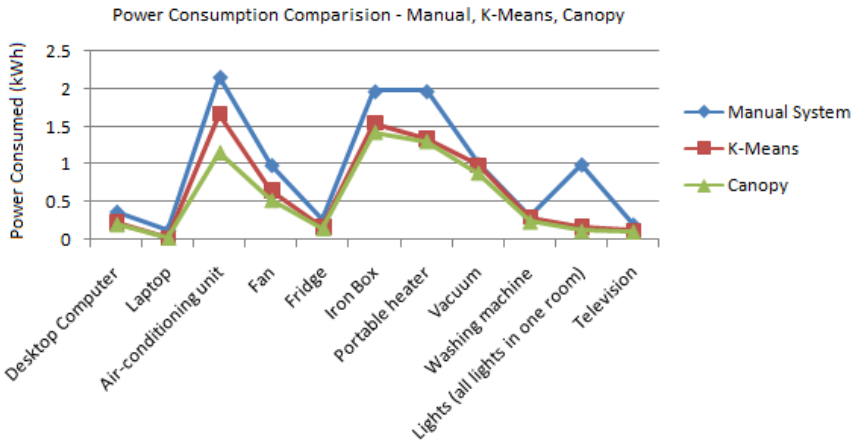

Figure 1: Various levels of energy consumption of home appliances

The experimental result shows canopy clustering algorithm based power regulating system consumes less energy compared with remaining two approaches K-means and manual based clustering algorithms approaches. The X-axis of this graph shown home appliances and Y-axis of this graph denotes various energy consumption levels.

6. CONCLUSION

In this research work, energy consumption level in home automation system is compared by implementing canopy algorithm, k-means algorithm and manual system. Based on the analysis this research work proved that canopy algorithm based home automation system performs best compared to other two methods such as k-means algorithm based home automation and manual method. Manual system consumes more energy, canopy algorithm consumes less energy and k-means algorithm consumes lesser than manual and higher than canopy.

REFERENCES

1. Isaac Machorro-Cano, Giner Alor-Hernández et al., "HEMS-IoT: A Big Data and Machine Learning-Based Smart Home System for Energy Saving", MDPI, Energies 2020, 13, 1097; doi:10.3390/en13051097.
2. Min Li, Wenbin Gu et al., "Smart Home : Architecture, Technologies and Systems", ELSEVIER, Procedia Computer Science 131 (2018) 393–400, ScienceDirect.
3. Dhiren Tejani, Ali Mohammed A. H. Al-Kuwari et al., "Energy Conservation in a Smart Home", ResearchGate, May 2011, pp.1-6.
4. Adela Bara, Bogdan Tudorica et al., "Solutions for Big Data Processing and Analytics in Context of Smart Homes", Solutions for Big Data Processing and Analytics in Context of Smart Homes, Database Systems Journal, vol. IX/2018, pp.80-89.
5. Shen Bin, Liu Yuan et al., "Research on Data Mining Models for the Internet of Things", 978-1-4244-5555-3/10/$26.00 ©2010 IEEE.
6. Akshat Savaliya, Aakash Bhatia et al., "Application of Data Mining Techniques in IoT: A Short Review", International Journal of Scientific Research in Science, Engineering and Technology, 2018 IJSRSET, Volume 4, Issue 2, Print ISSN: 2395-1990, Online ISSN : 2394-4099.
7. Ali Abdul-hussian Hassan, WahidahMd Shah et al., "Clustering Approach in Wireless Sensor Networks Based on K-means: limitations and Recommendations", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019..
8. K. Sathishkumar, E. Balamurugan, [2019], "Efficient Clustering Methods and Statistical Approaches For Gene Expression Data", Journal of Advanced Research in Dynamical and Control Systems, ISSN 1943-023X. Vol 11, Pages 440-447.
9. Sathishkumar. K, V. Thiagarasu, Dr. Balamurugan E, David Otto Arthur , "Clustering Mutual Outline for Multi Assessment Temporal Data and cancer Data", International Journal of Computer Sciences and Engineering, Vol.6, Issue.1, pp.75-83, 2018.
10. Sujith Jayaprakash, E. Balamurugan, “A Comprehensive Survey on Data Preprocessing Methods in Web Usage Mining”, International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, pp. 3170-3174.
11. K. Sathishkumar, Dr. V. Thiyagarajan E. Balamurugan, “Clustering Techniques for DNA Micro Array Based Gene Expression Data Survey” , International Journal of Engineering Sciences & Research Technology (IJESRT) ISSN: 2277-9655 June 05,2015.