International Journal of Pure and Applied Mathematics
Volume 119 No. 14 2018, 67-74
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Adaptive Clustering Strategy for heterogeneous and dynamic data for IoT scenario

Priya Dogra[1], Rakesh Dhiman[2]

[1]M.Tech(Scholar), [2]Assistant Professor, Dept. of CSE, NITTTR, Chandigarh
**Corresponding author and email:*
**[1]priyadogra.cse@nitttrchd.ac.in*
**[2]raakeshdhiman@gmail.com*

Abstract

Now a day's IoT is a revolutionary vision pertaining to the fact that everything could be connected to the Internet. With the increase in number of Internet users, IoT clients are also increasing. Data generated by devices is enormous and becomes troublesome to handle it for efficient storage and analysis. So, we need an efficient approach to manage multi-dimensional IoT data. IoT data can only be extracted by some appropriate data mining algorithms and transformed into an understandable form. To generate such relevant information, majorly clustering techniques are employed in data mining. Clustering basically helps in aggregation and organization of data in a structured form. Hence, in proposed work, we focus on developing various clustering methods pertaining to IoT data collected in different scenarios such as smart homes, intelligent traffic, smart weather information system etc. Data streams in different scenarios have different type and patterns for which no single clustering technique is optimal. Thus, IoT data streams are analyzed and accordingly appropriate clustering technique is applied. Conceivable result of this proposed approach is: Guidelines that helps in deciding suitable clustering scheme for the given IoT data set which results in optimal performance.

Keywords- *IoT, Data mining techniques, Data streams, Data Stream Clustering.*

1. Introduction

Internet of things

IoT in the coming times is going to become a fundamental concept in Internet technology because of its innovative vision of connecting devices and their communication via internet. Its rapid growth increases the risk of managing massive data. In IoT, the data is captured from different devices such as sensors, Radio–frequency Identification (RFID), ZigBee etc. Extracting specific information from
Such a wide-ranging data ware house is difficult. To handle this issue, data mining approach is employed in the proposed work.

1.1 Data mining

According to the definition of Data mining it includes three steps:

Data preparation – This means the preparation of data for mining. It includes three sub steps:

- Firstly integrate the data from various data sources then clean noise from data.
- Extract the data
- Preprocess the data

The outcome of the data preparation phase is the final data set. Once available data sources are identified, they need to be selected, cleaned, constructed and formatted into the desired form.

Data mining –Data mining finds valuable information hidden in large volumes of data. Data mining is the analysis of data and the use of software techniques for finding patterns and regularities in sets of data. After finding patterns it generates relationship within the data.

Data presentation- It is the presentation of the mined data to the users.

Fig. 1.1 Overview of Data Mining[17]

1.2 Data Mining techniques

The various data mining techniques which are used to extract the data are:-

- Clustering
- Classification
- Pattern analysis

Clustering

It is a procedure of partitioning informational indexes into subsets

on the basis of uniformity. It is an unsupervised learning process. It partitions information into groups these groups are known as clusters. Clustering has assumed a crucial part in assortment of fields-brain science and other sociologies science, measurements, design acknowledgment, data recovery, machine learning and information mining.Clustering is a data mining procedure that makes significant clusters which have comparable qualities. It characterizes the classes, while in the arrangement strategies clusters are doled out into predefined classes. Let us have a case to make the clear idea of clustering- We can take book administration in the library for instance. In a library, there is an extensive variety of books on different themes accessible. The test is the means by which to keep those books in a way that per users can take a few books on a specific point without bother. By utilizing the clustering system, we can keep books that have a few sorts of similitude in one group or one retire and name it with an important name. In the event that per users need to snatch books in that point, they would just need to go to that rack as opposed to searching for the whole library.

Classification

Classification is a data mining technique which requires machine learning. Essentially, it is utilized to arrange information into one of a predefined set of classes or groups. Grouping strategy makes utilization of scientific strategies, for example, choice trees, straight programming, neural system and measurements. In classification, we build up the product that can figure out how to arrange the information things into gatherings. For instance, we can apply classification in the application that "given all records of representatives who left the organization; anticipate who will presumably leave the organization in a future period." For this situation, we separate the records of workers into two gatherings that named "leave" and "remain". And after that we can ask our data mining software to order the workers into particular gatherings.

Pattern analysis

Pattern mining is a data mining technique which deals with identifying relevant patterns between data sets where the values are delivered in a sequence. It is usually assumed that the values are discrete, and thus time series mining is closely related, but usually considered a different activity.

1.3 Applications of IoT Smart Home

Now a day it is one of the most popular techniques of IoT, because of its availability and reliability. From no. of products in the market are now controlled by the users with their voices and senses. This makes the environment more lively and connected.

Wearable

Watches are not only use for just time. The most revolutionary product is Apple watches which have features like calling, text messaging and many more. They also give GPS locations to the users.

Smart Cities

IoT make the cities as Smart cities where each and every device has connected to each other. This makes the cities smarter than ever. With the proper and reliable connection we can solve the traffic issues and minimize the noise crime and pollution.

Connected Cars

In IoT the vehicles are interconnected to each other through internet. They can access each other data in a connected environment. This will help the people to make a connected and safer environment. One can easily access the others location.

1.4 Challenges and open Research Issues in IoT

The various challenges in IoT are:

Security

Security is the essential pillar of IoT. IoT is developed on thousands of SNs and massive data is generated through communication system, so the data and its communication need to be secure. No one should be able to steal data easily.

Privacy

In most deployments, the data generated must be protected from unauthorized access. No one should be able to decipher the data in static form or while in communication. A Federal Trade Commission report entitled "Internet of Things: Privacy & security in a Connected World" found that fewer than 10,000 households can generate 150 million discrete data points every day. This creates more entry points for hackers and leaves sensitive information vulnerable.

Lack of standards

In IoT, large numbers of heterogeneous devices are networked so it does not have a single platform for standardization numerous varying data storage and communication technologies /standards are used by these devices. Hence, to draw a common standard for such a scenario is extremely difficult. Further, to monitor and enforce such a standard if one exists is even more challenging.

Data management

The data generated through the IoT environment is huge and dynamic in nature. So, to extract useful data and manage it in such a big environment is too difficult. To handle this dynamism coupled with heterogeneity of data formats is a real issue. Data is the base of any IoT industry, so it has to be safe, secure and managed.

Multiple Connectivity

In IoT, each and every device has to be networked. So, in real IoT environments 100’s of devices are connected together, each having its own communication standard. Heterogeneity in communication protocols and interfaces make it extremely troublesome to provide seamless connectivity.

2. Related Work

Data Stream ClusteringDaniel Puschmann *et al.* [1] proposed a clustering scheme for dynamic IoT data streams. In this research paper, he focused on the state of the art and discusses the benefits and drawbacks of different stream cluster algorithms. He analyzed stream data with concept and data drifts. He introduced the concept of adaptive clustering method which automatically computes the best clusters based on the data distribution. He compared the method against synthesized data set and chose a silhouette coefficient to evaluate the experimental results. He gave various clustering techniques namely –
K mean
K-mean++ DBSCAN

Praveen Kumar [12] proposed a bunching plan to deal with enormous measure of information in IoT. The thing that continues going by in his brain is the tremendous measure of IoT applications. There he produce an information stream bunching plan which is unique in relation to the conventional information as just a single output is feasible for mining in light of its dynamic handling and also the stream keeps on developing. The enormous information stream created by IoT applications are considered of incredible significance and for removing data, information mining calculations can be connected to the produced information stream. Information stream bunching one of the critical procedures of information mining is very useful in grouping the comparable information protests and identifying exceptions. The target of his examination is to survey the diverse information stream grouping calculations that are utilized to perform bunching on the information stream created by IoT applications.

He utilized two bunching plans –

K-medoid
DBSCAN

Narendra Sharma *et al.* [22] concentrated on the information mining calculations which are then executed in WEKA instrument. The fundamental point of this paper is to give an itemized presentation of WEKA grouping calculations. WEKA is the information mining apparatus. It is the most straight forward device for characterize the information in different sorts. It is the main model for give the graphical UI of the client. As per them, Data mining programming is one of various logical apparatuses for breaking down information. It enables clients to examine information from a wide range of measurements or edges, arrange it, and outline the connections recognized. WEKA is an information mining apparatuses. It is contain the many machine inclining calculations. It is give the office to arrange our information through different calculations. In this paper we are concentrate the different grouping calculations. Clustering or grouping is the task of appointing an arrangement of items into groups) so that the articles in a similar group are more like each other than to those in different bunches. Our primary mean to demonstrate the correlation of the diverse distinctive bunching calculations of WEKA and discover which calculation will be most appropriate for the clients.

Data mining in IoT Rieko *et al.* [16] proposed a hole between distributed computing and big data in the transformative procedure of IoT. By checking and systematizing the procedure beginning from big data, not withstanding an investigation from network registering to distributed computing? Does the example of idea advancement progress toward becoming clearer? The reason for their review is to recognize the development of an idea room an innovation to an administration by watching the IoT watchword move on account of Big information. They have utilized a procedure to concentrate learning with time scale data from huge measures of archive information by viably utilizing content mining

Feng Chen *et al.* [17] they concentrated on the idea of information mining in IoT. Information mining includes finding novel, intriguing, and conceivably valuable examples from vast informational collections and applying calculations to the extraction of concealed data. The target of any information mining procedure is to construct a proficient prescient or distinct model of a lot of information that best fits or clarifies it, as well as ready to sum up to new information. They have applied the pattern mining technique to extract the information from a huge data but it was not that much efficient because it is difficult to find the pattern in such a huge environment. The Internet of Things idea emerges from the need to oversee, robotize, and investigate all gadgets, instruments, and sensors on the planet. Information mining advances are coordinated with IoT advances for basic leadership support and framework streamlining.

Shen Bin *et al* [20] proposed four information digging models for the Internet of Things, which are multi-layer information mining model, dispersed information mining model, Grid based information mining model and information mining model from multi-innovation combination viewpoint. Appropriated information mining model can take care of issues from saving information at various destinations. The Internet of Things will create expansive volumes of information. With respect to the few key issues in information mining of IoT are likewise talked about in their research papers. As indicated by them, improvement heading of the up and coming era of Internet, the Internet of Things draws in numerous considerations by industry world and scholastic circles. IoT data has numerous attributes, for example, disseminated capacity, mass time-related and position-related information, and restricted assets of hubs and so forth. This makes the issue of data mining in IoT.

Yan Chen *et al.* [21] now a day, an ever increasing number of individuals wind up plainly affectionate in the genuine and advanced universes, and constantly in cooperative collaboration. In this circumstance, to ensure the online clients, for example, the usernames and passwords is essential. As indicated by him one savvy approach in view of information mining called Associative Classification appears a potential arrangement that may successfully distinguish phishing sites with high exactness. They initially survey connection between the information mining and the Internet of things. At that point they highlight the potential arrangement. At long last, they close the potential arrangement is a successful approach to secure the online clients. They said that

Internet social networking and online transactions daily confronted with the most important question is harmful phishing sites. In real life, this is because the fishing site is the network security issues cannot be ignored. In the near future, we consider the function-based content. Their objective is to improve the collection of functions. They believe this will be potential research directions. The reason is that it helps to understand the attacker's behavior, which may help us to improve the performance of the method. Finally, use data mining process and data mining cases, and by studying data mining method and found to be the selected method is effective, especially to protect online users.

IoT & Big Data Hongming *et al*. [6] proposed a framework that identifies the need of data mining in IoT, its processing and mining areas of IoT big data. They categorized IoT data in cloud platforms where several conflicts are generated related to infrastructure resources, multitenant storage etc. They have also discussed the issues to implement IoT with cloud. During the research they faced the challenge of handling these massive heterogeneous data in highly distributed environments, especially in cloud environments. They also give a framework of IoT in which data-processing process is given to show the overview of related studies on the view of application. They, also discuss several open issues related to recent IoT developments and applications. Based on their framework of IoT data processing process, this research was divided in four modules. He also generated experimental results in which the comparison of implemented algorithms have been identified.

Saral Nigam *et al*. [29] proposed the framework which takes a shot at performing information mining and pattern examination over deals, stock information which is gathered progressively by utilizing POS framework executed utilizing RFID labels connected to products. According to them ,The IoT licenses detecting of the items and remotely controlled access crosswise over different system foundations, building up open doors for a superior combination between this present reality and mechanized world. Their principle concentrate is on the IoT applications, bringing about upgraded proficiency, precision and better financial outcomes. Each protest can be recognized interestingly by the utilization of its implanted registering framework however these items can interoperate inside the current foundation of the Internet.

Jeu Young Kim *et al*. [19] proposed the home space which is the most critical research territory of IoT. Presently gigantic information is produced by home brilliant gadget. There is developing concern, yet the past works didn't deliver enough to oversee and investigate home information. The motivation behind this review is to portray and look at to deal with the collected home IoT information in view of SWO and SWO examination stage. They demonstrated the execution of SWO investigation stage and a contextual investigation utilizing genuine information from brilliant metering gadgets for examination of machine use designs. This paper proposed not just the home IoT administration model and strategies but additionally give a model to driven home information mining strategy. They characterize content assets for managing the stream information.

Decebal *et al*. [27] proposed a smart grid approach, which consolidates building energy flexibility in real time environment by adding energy disaggregation in buildings. They have discussed the problem formulation and evaluate their results with two model machines i.e. Figured Four-Way Conditional Restricted Boltzmann Machines (FFW-CRBMs) and Disjunctive FFW-CRBM. In this paper, they have proposed an IoT structure to perform at the same time and progressively adaptability ID and forecast, by making utilization of Factored Four Way Conditional Restricted Boltzmann Machines and their Disjunctive adaptation.

Wen-Tsai *et al*. [28] gives a cutting edge writing audit on monetary investigation and evaluating models for information accumulation and remote correspondence in IoT. Remote sensor systems (WSNs) are the primary segments of IoT which gather information from the earth and transmit the information to the sink hubs. Moreover, with billion of gadgets associating with the Internet, it prompts many difficulties in effectively controlling and dealing with IoT's sensors. Thus, new methodologies with higher productivity and greater adaptability to adjust to dynamic IoT systems should be created. This paper has given a complete review of the financial and valuing hypothesis and additionally their applications in information accumulation and correspondence of IoT. They have talked about some open issues in IoT.

Antonio M. Ortiz *et al*. [15] proposed a framework i.e. SIoT Architecture. According to them, with the increase of IoT users, no. of devices used in IoT environemt are also increasing and with this the data generated by these sensors has also increased. To handle SN, IoT data, devices was itself a big task. So they have generated a functional idea which created a group between Internet of Things (IoT) and SNs. In this structure, the data originating from IoT, and the SN conveys it to enable human-to-device communications. This paper investigates the novel approach to handle IoT data, indicated by SIoT. Hence, this paper initially addresses an entire view on SIoT and key viewpoints to imagine the genuine omnipresent figuring. They emphasized on role of SN and also new challenges and open issues in SIoT.

Akshay Kumar *et al.* [26] He examined a significant issue in incorporating distributed computing and IoT. He has generated a cloud based IoT framework in which the IoT devices attached in IoT cloud. They have proposed Greedy Heuristic algorithm to formulate their research. Additionally they proposed the consolidated issue of transferring IoT information from an arrangement of sensor gateways and effective replication of this information on conveyed distributed storage. They investigated the integrity in cloud computing and IoT.

M. Marjani *et al.* [23] they gave an overview of IoT & Big data in which the comparison between IoT & big data as well as similarity between the two has discussed. They have proposed four analytical systems to analyze the IoT data efficiently. He investigates the various problems generated with the applied approach and gave an IoT & big data analytic model to verify its research. They implemented their research by using notable test cases. He examined the various challenges and open research issues related to IoT & Big data analytics.

2.1 Inferences drawn from literature review

After the critical review, the following inferences are drawn:

- Most of the authors have proposed clustering algorithms to mine the data in IoT environment.
- K-medoid algorithm takes more time for the execution while random clustering is the fastest one as comparison to all other data clustering algorithms.
- Because of rapid growth in IoT world, it is too difficult to handle the data streams so more robust algorithms will be needed to perform clustering.
- Too much data – IoT data is itself a largest type of data on the Internet. It is difficult to manage such a huge data in a IoT environment.
- Data is not secure and safe.
- According to Gartner's research, processing large quantities of IoT data in real time will increase workload on data centre. Service providers face
- difficulty of privacy, security and reliability of data as well as devices.
- In IoT world tones of data is floating around, so there is always a risk of data to be stolen.
- The growth of IoT has led to the increase in the huge volume of real world streaming data.

3. Problem Formulation and Proposed Work

Based upon the related work, following is the overall statement of the problem:

"To develop metrics that help in deciding a scenario specific clustering technique to efficiently manage (storage/ retrieval) information for a given IoT scenario"

3.1 Objectives

To meet above objective, following sub-objectives are proposed.

- To develop metrics that help in deciding clustering method to be adopted for a specific scenario.
- To devise a method that dynamically gives comparative study on Adaptive clustering by using clustering techniques i.e. k medoid clustering.
- To study Adaptive cluster with feature extraction.
- To perform simulation experimentation to conclude and verify clustering strategy.

3.2 Assumptions

The following are the assumptions of this research work:

- Data is dynamic and multidimensional i.e. heterogeneous in nature.
- Too much workload on the data centers.
- Data from sensors are always raw and low- level data.
- Because of huge amount of data to handle, so there must be an efficient and effective way to process it.
- The aim of data mining is to achieve a comprehensive data analysis for end users.

3.3 Methodology Flowchart

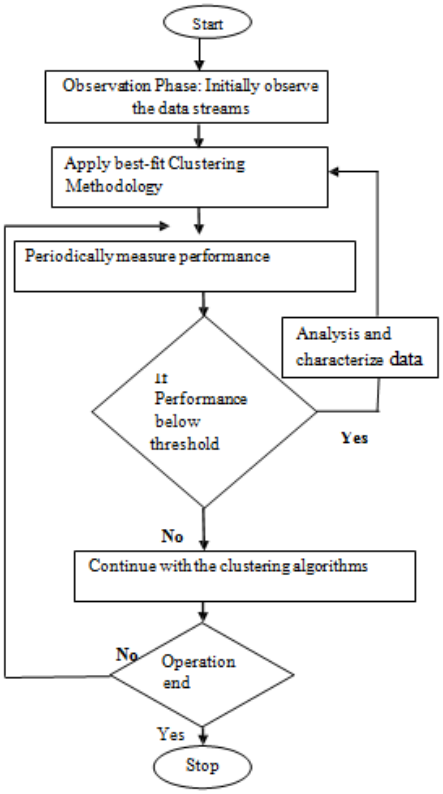

Fig. 3.1 Methodology Flowchart

Methodology adopted to carry out our proposed work is shown in following flowchart: fig 3.1

3.4 Methodology of proposed Adaptive Clustering Scheme

Step1: Initially the data streams are analyzed collects from (SN and CN).
Step 2: Apply the best fit clustering algorithm to mine the data.
Step 3: Initially the clustering algorithm measures the performance periodically. If the performance is below threshold it again goes to Step 2 and again analyzes & characterizes the data streams.

 Step 4: Continue with the clustering algorithms to generate the clusters of data.
 Step 5: Stop

3.5 Tool Used

MATLAB stands for MATrix Laboratory. It is a programming language designed by MATHSWORKS. It initialized out as a matrix language where linear-algebra programming language was easy. MATLAB could execute both under inter-active sessions and a batch task. MATLAB is implemented to offer students effortlessness in programming language. The issue based on MATLAB illustrations has been offered in easy and simple path to create the learning speed faster and effective.

In this MATrix Laboratory is 4^{th} generation high level programming language and effective manner structure for integral calculation and visualization?

It defines MATrix changes, plotting of methods and data designing of methods, creation of consumer interface, interfacing with written in some other languages like C++, C, JAVA and FORTAN, study of the data, designed algorithm and make structures and applications.

Simulation tool MATLAB 2016a is utilized in every surface of calculation mathematics. Subsequent are some normally used math's calculations where it is utilized most generally:

- Matrices and Array deals.
- 2 and 3 dimensional plotting and graphs.
- Linear Algebra
- Algebraic Equ. And Non-linear methods
- Statics and data Analysis
- DI (Differential eq.) and Calculus
- Numeric Evaluations
- Integral
- Transforms and Curve Fitting
- Several other Normal Methods

3.6 Characteristic Of Matlab

The basic characteristics of MATLAB are following:

- MATLAB is a HLL (High level Language) for integral calculation, visualization and applications design.
- MATLAB also gives an inter-active structure for iterative examination, model and issue solving.
- It gives a famous library of mathematic methods for Statistics, Numerical, filtrating and optimization and resolving ODI (ordinary Differential Eq.).
- It offers pre-defined a graphics for visualizing information and tool for generating custom plot of the diagram.

4. Reason for selection

- It provides a better interface to the user than the other data mining tools.
- In MATLAB one can easily work without having the deep knowledge of data mining techniques.
- MATLAB provides GUI interface so it is easy to access.

5. Conclusion and Expected Outcome

Proposed work presents an efficient approach for managing huge IoT data streams using clustering techniques. It gives the comparative study on various clustering techniques. In the detailed literature survey mentioned above, we present a deep and thorough study in reference to IoT data, data mining techniques and clustering algorithms. IoT data being dynamic and multidimensional in nature creates lot of workload on data centre. Thus, data mining techniques are employed in the same for managing this type of overhead. Moreover, clustering methods help in filtering out useful information from IoT data. Data generated in a scenario may change over time in respect of sampling rate, size and staleness and hence no single clustering strategy will be suitable for every situation, so, proposed work thrives to apply adaptive clustering strategy where clustering scheme applied for a dataset varies accordingly.

5.1 Expected Outcome

Our main focus is to apply suitable clustering algorithm on situation-specific IoT data. This proposed strategy is supposed to reduce the complexity of IoT data streams collected in data centre, time factor will also reduce Moreover it improves the accuracy and overcomes the error. Useful information can also be inferred from the given IoT data using data mining schemes. Further this approach helps in eliminating replicated data streams and efficiently using the storage capacity of data centre.

References

[1] Daniel Puschmann, Payam Barnaghi and rahim Tafazolli, “Adaptive Clustering for IoT data streams” IEEE Internet Of things journal, Vol. – 4, Issue 1, pages 64 - 74, February 2017.

[2] Geng Yang, Li Xie, Matti Mäntysalo, Xiaolin Zhou, Zhibo Pang, Li Da Xu, Sharon Kao- Walter, Qiang Chen, and Li-Rong Zheng, “A Health- IoT Platform Based on the Integration of Intelligent Packaging, Unobtrusive Bio-Sensor and Intelligent Medicine Box” IEEE Transactions on Industrial Informatics, Vol.10, Issue 4, pages 2180 - 2191, Nov.2014.

[3] Hongming Cai, Li Da Xu, Boyi Xu, Cheng Xie, Shaojun Qin, Lihong Jiang, "IoT-Based Configurable Information Service Platform for Product Lifecycle Management," Industrial Informatics, IEEE Transactions On Industrial Informatics, VOL. 10, Issue 2, pages- 1558 - 1567, May 2014.

[4] M. Ester, H. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in Proceedings of the 2nd International Conference on Knowledge Discovery and Data

Mining , Vol. 4, pages 226-231 , August 2016.
[5] J. Agrawal, S. Soni, S. Sharma, and S.Agrawal, "Modification of density based spatial clustering algorithm for large database using naive's bayes' theorem," in Proceedings of the 4th International Conference on Communication Systems and Network Technologies, Vol.1, pages 419 – 423, April 2014.
[6] Praveen Kumar, "Data Stream Clustering in Internet of Things" SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE)- volume 3 Issue 8, pages 103-108, August 2016
[7] Chun –Wei Tsai, Chin – Feng Lai, Ming-Chao Chiang and Laurence T. Yang, "Data Mining for Internet Of Things: Survey", IEEE Communications Surveys & Tutorials, Vol.16, Issue1, pages 77 - 97, 2014.
[8] Andrela Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista and Michle Zrorzi , "Internet of things for smart Cities" IEEE Internet of Things Journal, Vol.1, No.1, pages: 22 – 32, February 2014.
[9] Antonio M. Ortiz, Dina Hussein, Soochang Park, Son N. Han and Noel Crespi "The cluster Between Internet of Things and Social Networks: Review and Research Challenges", IEEE Internet of Things Journal, Vol.1, Issue 3, pages 206 – 215, June 2014.
[10] Rieko Kataoka, Naoshi Uchihira, Yasuo Ikawa, "The Evolutionary Process of IT Concept Words: A case study on Bigdata", Proceedings of PICMET 16; Technology for Social Innovation, Vol.2, pages 1983 - 1992 March 2016.
[11] Feng Chen, Pan Deng, Jiafu Wan, Daqiang Zhang, Athanasios V. Vasilakos and Xiaohui Rong "Data mining for the Internet Of things: Literature review and challenges ", International Journal of Distributed Sensor Networks , Article No. 431047 Vol.11 ,Issue 8, Aug. 2015.
[12] Jeu Young Kim, Hark-Jin Lee, Ji Yeon Son and Jun-Hee Park "Smart Home Web Objects–based IoT management Model and Methods For Home data mining" IEICE APNOMS, pages 327 - 331, Sep.2015.
[13] Shen Bin, Liu Yuan, Wang Xiaoyi "Research on Data Mining Models for the Internet" 10th International conference on Image processing IEEE, pages 127-132, April 2010.
[14] Yan Chen, Ai-xi Han Cai-Hua Zhang, "Research on data Mining Model in IoT" International conference on Automation, Mechanical Control and Computational Engg. AMCCE, April 2015.
[15] Narendra Sharma, Aman Bajpai, Mr. Ratnesh Litoriya "Comparison the various clustering algorithms of weka tools "International Journal of Emerging Technology and Advanced Engineering, Vol.2, Issue 5, pages 73-80, May 2012.
[16] Mohsen Marjani, Fariza Nasaruddin, Abdullah Gani, Ahmad Karim, Ibrahim Abaker Targio Hashemi, Aisha Siddiqa and Ibrar Yaqoob "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges" Vol.5, pages 5247 – 5261, March 2017.
[17] Akshay Kumar, Nanjangud C. Narendra , Umesh Bellur "Uploading and replicating IoT data on Distributed Cloud storage" IEEE 9th Conference on Cloud Computing, pages 670 - 677, April 2016.
[18] Deccebal Constantin Mocanu, Phuong H. Nguyen, Madeleine Gibescu and Antonio Liottta "Big IoT data mining for real time energy disaggregation in buildings" IEEE International Conference on Systems, man and Cybernetics, Vol.4, No.3, pages 003765 – 003769, Oct. 2016.
[19] Saral Nigam, Shika Asthana, Punit Gupta, "IoT based Intelligent Billboard using Data Mining" Ist International conference on Innovation and Challenges in cyber Security ICICCS, pages 107 – 110, August 2016.