Research on Data Mining Models for the Internet of Things

Aakanksha[1], R. Sai Varsha[2], Ashima Mehta[3]

UG Students, Department of Computer Science Engineering[1,2]

Head of Department, Department of Computer Science Engineering[3]

Dronacharya College of Engineering, Gurugram, Haryana, India

Abstract: *Data mining is a technique or method which is used to extract datasets from different source and the hidden part of applications. Data mining is also used to make a pattern form a large amount of data by which the data will be more accurate and visible to the user. Data mining is used to enable the user to get a summary view of data according totheir research. This paper discusses three data mining model of IOT that how data mining model helps IOT in growing the business and how to improve IOT using data mining technique. The future scope of IOT is discussed.*

Keywords: Internet of Things, Data mining models, clustering algorithm, automatic identification

I. INTRODUCTION

In 1999 the term Internet of things was introduced by Kevin Ashton while working in Auto-ID labs. IoT plays an important role in every spectrum of business, healthcare sectors and education sectors.

The Internet of things is a system of interrelated computing devices, mechanical and digital machines provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction. It is powerful tool that senses the object automatically and gives the expected output. The sensors and actuators are embedded with this and the device can work automatically.

An IOT device is connected to various software and servers to collect and exchange the data over the internet. IOT is an AI device that consumes a large volume of data so that data mining model for IOT is used to handle the data sets and select the accurate data then process it to the machine. Data mining has several clustering algorithms & the clustering algorithms is used to provide effective methods to analyse, manage and identify the data for the machine. Data mining in IOT is used to organize data in a particular order in a storage that will help the machine to retrieve data at any time.

II. RELATED WORK

This paper focuses on mining and managing the data using some suitable clustering algorithm and some works are done about query analysing about moving data which is generated from IOT. The author Jae-Gil Lee, at reference [12] provides a proper framework which is used to detect the moving datasets.

This paper also focuses on machine learning and knowledge discovery from sensor actuators that how to provide expected output form the knowledgebase. Here the Author Joydeep Ghos represents a framework that allows supervised learning under some condition and limitations.

The author Rohit Yadav Represents a proper data mining model of IOT that how various data mining model is used In IOT devices to store and organize the data in proper order that will be easy for user interface available in machine.

Azra Shamin suggested a framework for analysis of various datasets by using data mining. When combining the datasets, it forms heterogeneous data that are extracted from combined datasets and different sources of applications. Here the author Rumi Ghos represents a framework to combine multiple heterogeneous data to get the expected output from the IOT devices.

Amir Ahmad performs the clustering algorithm in various datasets to divide them according to their nature. The author represents a better characterization of cluster data. By applying clustering algorithm an IOT device can easily extract the required data from the knowledgebase.

Jayavardhana represents the future trends of IOT. The author also represents how IOT will be more helpful in the future and how data mining is used to extend the service of IOT in various fields.

In the new era of the Internet, the internet of things is also in the development stage. There are some pending works to be done about IOT. This paper represents a framework or scheme to aggregate some homogenous and heterogeneous data extracted from a different source and hidden part of the application

III. DATA MINING MODELS FOR INTERNET OF THINGS

Data Mining provides several clustering techniques to make the datasets more useful, avoid redundancy and present of Duplicate data in the database server or in knowledgebase.

After extracting information from a different source, Data is filtered and arranged in an effective way that enables a user to use the data. The data mining model is used to classify the data in a particular spectrum that will be easy to access the data. When data is filtered by the clustering technique, we can inject all the data to IOT device by which the device can work properly and will give the expected output.

Data mining model includes various techniques such as Data exploration, Data modelling, Model deployments. Data exploration is an algorithm used to extract data and find the pattern of datasets that helps to discover the knowledge. Data modelling is all about to give a proper shape to datasets and used to represent the data in way that will be easy to understand. Model deployment is the most fundamental job that is done by data mining model.

When extracting the data from different source, hence the extracted data sets may be homogenous and heterogeneous datasets. It is not possible to combine both homogenous and heterogeneous data. The author Rumi Ghos represents a framework that how to perform clustering operations in heterogeneous datasets.

After extraction of various datasets from different application, the datasets are needed to integrate the data and clean the noise of data.Every IOT devices have sensors that are connected and the sensors are collecting the data from different source then perform analytical operations on extracted datasets to make them relevant.

According to the data mining architecture of IOT, Data mining model for IOT is divided into three different types of models. Such as: -

- Multi-Layer data mining model for IOT.
- Distributed data mining model for IOT.
- Grid based data mining model for IOT.

Multi-Layer data mining model for IOT

Now days, IOT devices are producing and collecting a large volume of data from different sources. Data streams are continuously generated from internet of things. This layer is used to manage the collected data then saves it to the data warehouse for future purpose.
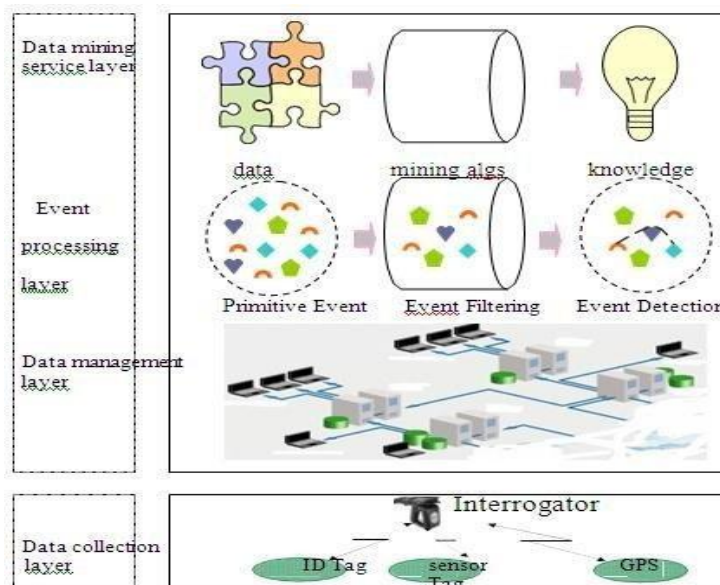

Fig-1 multi-Layer data mining model for IOT

In the (fig-1), multi-layer data mining model for IOT has also four different types of layers: data collection layer, data management layer, event processing layer and data mining service layer.

Data collection layer is used to collect data from different application server. In the data collection layer, there may be the present of duplicate data, fault tolerance and miss-match of datasets. This problem is solved by data collection Layer.

Data management layer is used to manage the data that are collected or extracted from application server. After collection of data, there may be a present of homogenous and heterogeneous data, Hence Data management layer is used to differentiate between homogenous and heterogeneous datasets and manage them in an effective way. Data management layer is used to store the data in an active online catalog or in data warehouse for future use.

Event processing layer includes event cleaning, event filtering and event detection. Here event means datasets. Event processing layer is used to clean the data and ensure that there is no bug and the data is various free. In this layer datasets are filtered and detects the datasets if the layer finds any duplicate data. By using event processing layer IOT devices can execute event-based query and solve complex problems in an effective manner.

Data mining service layer is a combination of data management layer and event processing layer. In this layer various operations of clustering takes place to make the whole data sets accurate and visible for user interface. We can predict, forecast, analyse the various patterns. It provides service to user-interface in which user can interact. By using data mining service layer, the duplicate and redundant datasets can be detected by IOT devices.

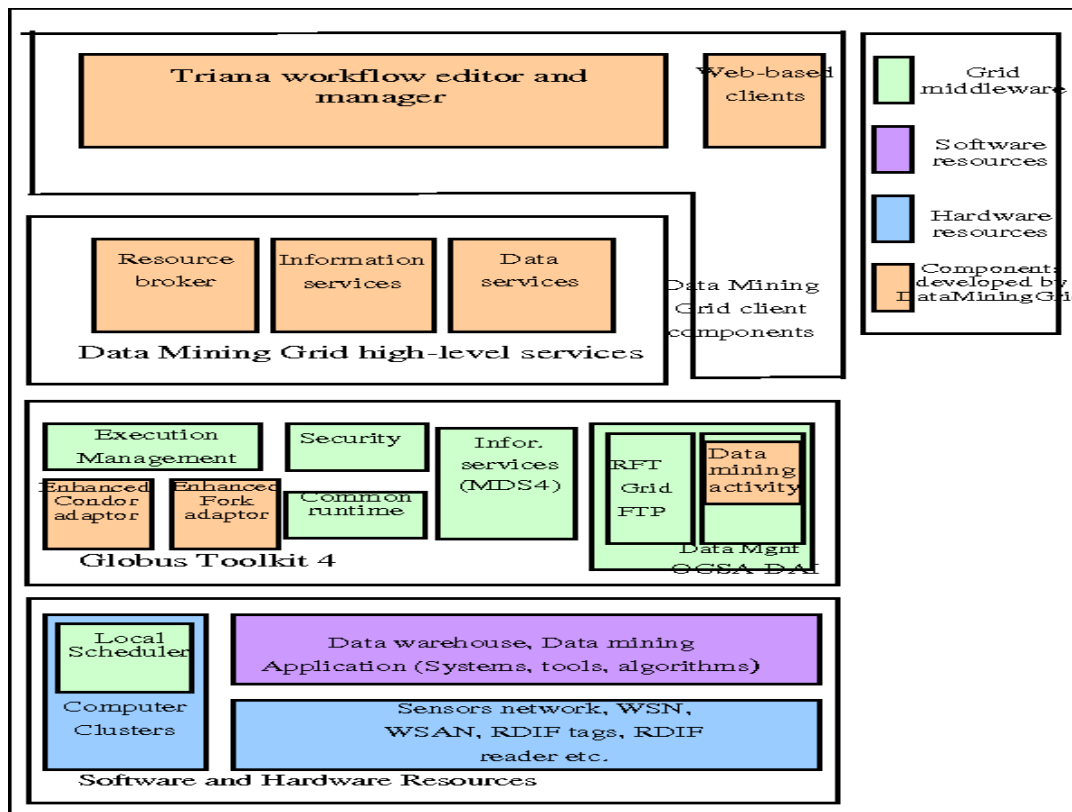**Distributed data mining model for IOT**

Fig-2 Distributed data mining model for IOT

In the (Fig-2), Distributed data mining model for IOT is used to deal with the data coming from different application server. It has two types' nodes such as Local node and Global node. Local node consists of management of database operations and Global node is used to mine the data.

It is also used to solve complex problems and process that raw material to provide the necessary output according to the user's requirement. Basically, it is used to manage the data and resources available in

distributed database. Distribute data mining model contains many nodes in it, some of them local node and some are global node. This data mining model distributes the datasets in different nodes and enables the nodes to do the frequent operations on various datasets. The local node and global node can communicate with each other. The whole operation phase in distributed data mining model takes place in Global node then after operation the datasets are stored in local node warehouse for proper management and integration of datasets.
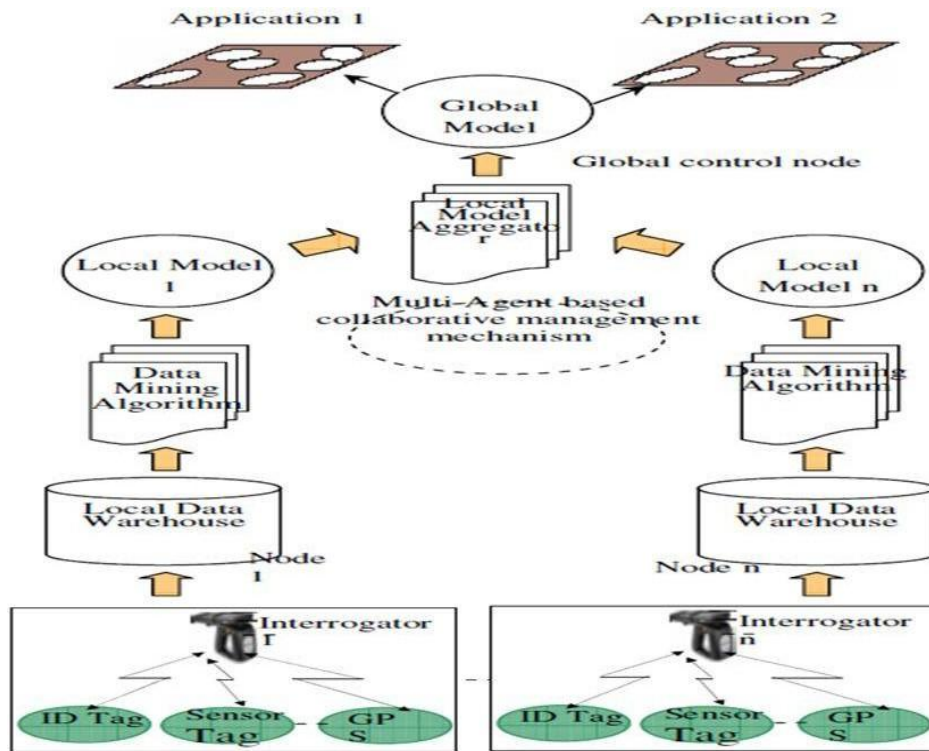
Grid based Data Mining Model for IOT

Fig-3 Grid based data mining model for IOT

In (Fig-3), Grid based data mining model for IOT is used to provide computational functionality in heterogeneous environment. This model follows non synchronized method that means at a time more than one task can be executed parallel, so that the waiting time of data mining is reduced and performance increased.

The basic aim of IOT is to connect the physical devices to internet to make them smart by which the physical devices can work properly when they are connected to a network.

According codd's rule the datasets must be stored to an active online catalog for easy access of database. P. Brezany et al[11]. Proposed a fundamental infrastructure called Grid Miner, which supports distributed online analytical processing and data mining that means the data generated by distributed data mining model must be stored to an active online catalog. Active online catalog is referred as data dictionary.

IV. KEY ISSUES IN DATA MINING MODEL FOR IOT

After extraction of datasets a suitable algorithm is used to clean the noise form data. Identifying the proper algorithm may be difficult to choose because some algorithm may be only for numerical datasets but we have to choose that algorithm which will be suitable to treat every dataset individually

Every data mining model follows synchronized method that means it supports parallel execution. In parallel execution the IOT device should be designed in a way that every suitable algorithm can be applied to data objects efficiently.

IOT is a vast technology and if any organization wants to invest in it then the organization should have the proper knowledge about IOT that is why understanding IOT may be difficult for some user.

Another key issue is security and privacy. IOT devices are dealing with massive quantity of data so that it may be difficult to provide security and privacy to every data objects.

In a database, every data is linked to each other that when removing the unwanted data objects from IOT, it may affect other data objects and creates confusion for user.

As IOT is growing rapidly, the device load increases here to handle to device load and management of data may be difficult.

IOT needs good secure internet connectivity; basically, organizations in rural area of INDIA may face connectivity issue. The network model for IOT must have the powerful tools that the model can detect any unwanted files easily. The network model should have the ability to solve complex problems easily.

After extraction of data sets, there may a present of both homogenous and heterogeneous datasets and we cannot combine homogenous data with heterogeneous data. So that may be a challenging issue for Every IOT devices.

Every organization needs a database engineer to implement database operation, query processing, data filtering and removing the duplicate data objects from knowledge base is not an easy task, we have to check every individual data objects. It is a time-consuming process for every organization.

In an IOT device, there may be many datasets are stored in it, to identify the address of datasets may be a difficult task to implement.

V. FUTURE SCOPE OF IOT

IOT is introduced in 1999 and from that day IOT is a demanded futuristic technology that can be used in order to make tasks easy. In every spectrum, IOT plays an important role in various industries like retail and manufacturing, healthcare, education.

- **Smart Home:** In this growing era of technology, IOT plays a vital role that we can implement IOT in Smart home also. Smart homes can be referred to as monitoring the conditions of home and control them over internet from anywhere and in anytime. There are some sensors embedded within smart home appliances so that by using sensors the smart home devices can sense the condition of home and do task automatically.
- **IOT Security:** IOT also plays an in an important role in Data security. By implementing IOT in security devices, It will be easy to track your data efficiently. IOT automatically monitors network and systems. It acts as a firewall between the client system and a server for threat detection and network visualization.
- Every IOT devices has millions of nodes and each node is provided a unique task and the ability by which they can automatically transfer data over the network.
- **Better data analysis:** As we all know, every organization needs better data analyser, but after IOT we don't need any data scientist for data analysis, IOT devices can analysis data automatically properly and accurately. IOT devices are acts as a decision-making system for all business organization. IOT provides an effective process to execute task simultaneously.
- **Healthcare:** In Health care sector IOT plays a vital role as it automatically monitors the health status of the patient. IOT devices act as an expert system for doctors that IOT devices can recommend the doctors which medicine is good for a particular patient or not.In hospitals IOT devices are used to recommend medicine related information. IOT devices act as expert system.
- **Education:** The future of IOT in education will enable new forms of teaching — such as fully-remote physical education and science courses or data-driven teaching methods. Educators can use this technology to inform methodologies and build lesson plans.
- **Cloud Computing:**
- In cloud, we can upload our sensitive information. As we know cloud computing is connected to the Internet can be harmful in many ways, and the use of spyware can get login and password information. By implementing IOT to cloud we can efficiently protect our data.
- **Big Data:** By implementing IOT to Big data, It will be easy to make decisions which data we should pick or not. We use IOT in Big data, to control the flow of data and select which data is suitable or not, and IOT also stores data byusing different storing Algorithms.

- **Smart City:** In future trends, IOT can play a vital role in developing smart cities. IOT will help in traffic management, Waste management control, Water resource management, housing issue, and pollution. IOT will help a pollutedcity like New Delhi, Gurgaon and Faridabad.
- **Industrial Automation:** In the field of industry, every industry wants to maximize their profit at minimum cost. IOT will help the industries in optimization that will save time, cost-efficient, minimum resource, and quality control.
- **Disaster management:** IOT can't stop disaster management from happening but IOT helps a lot when we are about to face any cyclone or earthquake, At that time IOT will predict, recovery, and efficiently respond to the situation

VI. CONCLUSION

Now a day IOT is growing technology and every organization are investing big in it. IOT helps organizations in decision making and in prediction of trends in coming days. This paper discusses data mining model for IOT that how data mining model will be more useful for IOT in growing. This paper also discusses about future scope of IOT and key issues in IOT. Although there are some issues but companies are investing big in it because IOT is cost effective.

Data mining model plays a vital role in prediction and discovery the knowledge in various domains. In IOT, Data mining is used to provide the selected output as per the user's requirements. IOT is in development stage and some more works to be done about IOT. In coming years, the demand of IOT will increase and it will be more efficient to use

REFERENCES

[1]. Rohit Yadav, Kapil Arora. "Data mining for the internet of Things: A survey. "Communications surveys & Tutorials, IEEE 16.1(2014):77-97.

[2]. Azra Shamim, Vimala Balakrishnan, Madiha Kazmi, and Zunaira Sattar, "Intelligent Data Mining in Autonomous Heterogeneous Distributed and Dynamic Data Sources", 2nd International Conference on Innovations in Engineering and Technology (ICCET'2014)

[3]. Rumi Gosh, Sitaram Asur, "Mining Information from Heterogeneous Sources: A Topic Modeling Approach".

[4]. Joydeep Ghosh. "A Probabilistic Framework for Mining Distributed Sensory Data under Data Sharing Constraints," First International Workshop on Knowledge Discovery from Sensor Data. 2007.

[5]. Amir Ahmad, Lipika De, "A clustering algorithm for mixed numeric and categorical data" Data & Knowledge Engineering Elsevier.

[6]. Gubbi, Jayavardhana, et al. "Internet of Things (IOT): A vision, architectural elements, and future directions." Future Generation.

[7]. Internet of things definition, available from<https://en.wikipedia.org/wiki/Internet_of_things >.

[8]. Multi-layer data mining model for IOT available from <https://www.researchgate.net/figure/Multi-layer-data-mining-model-for- IoT_fig2_321333161>

[9]. Distributed data mining model for IOT available from <https://www.semanticscholar.org/paper/Research-on-data-mining-models-for-the-internet- of-Bin-Yuan/822535c409890de3aae74b49b2bd8d4a59832fba>

[10]. Grid based data mining model for IOT available form<https://www.semanticscholar.org/paper/Research-on-data-mining-models-for-the- internet-of-Bin-Yuan/822535c409890de3aae74b49b2bd8d4a59832fbaa>

[11]. P. Brezany, I. Janice, and A. M. Tajo. "Grid Miner: a fundamental infrastructure for building intelligent Grid systems," Proc. 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05), IEEE press, 200, pp. 150~156.

[12]. Jae-Gil Lee, Jiawei Han, Xiaolei Li, Hector Gonzalez: "TraClass: trajectory classification using hierarchical region-based and trajectory-based clustering," PVLDB 1(1): 1081-1094 (2008)

[13]. "Data mining model for Internet of things" Research by Shen Bin , Liu Yuan , Wang Xiaoyi.

[14]. "A Research Direction on Data Mining with IOT" Reserch by purvi Prajapati, Jay Patel.