INTERNATIONAL JOURNAL OF SCIENTIFIC PROGRESS AND RESEARCH (IJSPR) ISSN: 2349-4689

Data mining for IOT: A survey

Rakshitha B.S[1], Nayana K.M[2], Pooja Darshini B.N[3], Divya B.M[4]
[123]*UG student,* [4]*Assistant professor,*
[1234] *Department of information science and technology, BGSIT*

Abstract-In this paper we have focused a variety of techniques, approaches and different areas of the research which are helpful and marked as the important field of data mining Technologies. The massive data generated by the Internet of Things (IOT) are considered of high business value, and data mining algorithms can be applied to IOT to extract hidden information from data. The term data mining is appropriately named as "Knowledge mining from data" or "Knowledge mining". Data mining techniques are used in many areas in the world to retrieve the useful knowledge from the very large amount of data.

Keywords: Data mining, Internet of things.

I.INTRODUCTION

The Internet of Things (IOT) and its technologies can be integrated through networks with networked instruments and devices and it is shown in theFig 2.1 IOT has been playing the main role in the world it was covers from traditional equipment to general household objects [1][2] and has been attracting the attention of researchers from academia, industry, and government in now a day. There is a super vision that all things in the world can be easily maintained and monitored, these things can be identified automatically by other, from this we can communicate with each other through internet in world wide range, and from this they can even make proper decisions by themselves. To do IOT smarter and better, lot of analysis in technologies are introduced into IOT. One of the most valuable

Technologyis Data mining. Data Mining is the process of exploration of data and analysis of data from the large quantity of data or by the automatic or semiautomatic means of large quantities of data to discover useful and meaningful patterns and rules.

Data mining have five major elements:

- Extract, transform, and load transaction data onto the data warehouse system.
- Data can be store and manage in a multidimensional database system.
- Provide data access to business analysts and information technology professionals.
- Analyze the data by application software.
- Present the data in a useful format, such as a graph or table.

Data mining involves discovering interesting things and potentially useful patterns from the large number of data sets and applying algorithms to the extraction of hidden information. They are many other terms are used for data mining, for example, knowledge discovery (mining) in data bases, knowledge extraction. The objective of any data mining process is to build an efficient predictive or descriptive model of a large amount of data that not only best fits or explains it, but it is also able to generalize to the new data. Based on a broad view of data mining functionality, data mining is the process of discovering interesting knowledge from large amounts of data can be stored in any of the form of data bases, data warehouses, or other information repositories.

Fig2.1 Internet of things

In remaining papers and the current paper contains four sections, as the roadmap is organized as follows. Section I begins with a brief introduction to the IOT and Data mining. Section II describes the data from IOT, explains about the data mining for IOT and the main purpose of an open issue of data mining. Section III is conclusion.

II. DATA FROM IOT

As everything on IOT consist of data with different kinds of information. There will be technical problems and issues on how to handle these data and how to bring out the information that has emerged. We use a simple taxonomy [3] to differentiate the types of data from IOT is to use "data about things" in which the description of data is given by themselves. In older days, the data was used to optimize the performance of the systems, infrastructure and things of IOT whereas nowadays the data results in interaction between humans, between human and systems, and systems that can provide services which is given by IOT.

Depending on these characteristics, a new type of data is emerged as "Big Data". Big data is no longer a business promotion. It's something real that we are facing in day to day life. It's known that the tools that are available for data analysis are not powerful enough to handle and analyses big data of IOT. We are still having a difficulty to put more than one Zettabyte into single storage system. The next issue that we are facing today is that the data cannot be processed by the available tool because the data from IOT is too big and too hard.

Since from many years there is a problem in handling and analyzing the large amount of data, there are many traditional methods which are efficient enough and that are implemented in the past those can be used to solve and handle the issues of big data problem. There are some methods found by previous data mining studies, those are random sampling, data condensation, divide and conquer and incremental learning. Among these, the best way to solve the big data problem of IOT is to have sensors. These sensors will include only the interesting data among all the data. The researchers have found that by using these sensors we can reduce the complexity of input data. In initial method, we use the principal component analysis (PCA) 0r in other word a dimension reduction method, which is used to reduce the number of features of input data. Recently, we use another trend called pattern reduction (PR). This PR is based on a different consideration. Comparing PR with PCA, this technology is aimed at reducing the number of patterns instead of reducing the number of dimensions during the convergence process. Thus, this method can be used to reduce the complexity of input data. Apart from these methods the big data issues are handled from newly emerged feature selection, distributed computing and cloud computing [4].

In forestall, organizations will have large amount of big data from various sectors like services, applications, and platforms they provide in the future. With the addition of handling these massive data, finding the hidden data had become the big task for the organizations because this task put them into the absolute position and leads in the production of innovative products. At last, data analysis for sensors and devices help us to develop a friendlier system

DATA MINING FOR IOT

Big data, knowledge discovery in database (KDD), data mining have the relationship in IOT. In this section, we will be discussing about the application mining technologies and a brief introduction of the data mining technologies for IOT by giving a simple example.

A. Basic idea of using data mining for IOT

Comparatively creating data is the easier way than analyzing data. These created data may explode and leads a serious problem in IOT. The researchers have given many solutions to solve the queries of big data in IOT [5]. We and all the systems are submerged by the unprecedented data without effective and efficient tools. When KDD is applied to IOT, from the viewpoint of hardware, the available solution for big data are cloud computing and relevant distributed technologies. In the condition of big data, it is almost certain that most KDD system available today and to process the large amount of data of IOT with the help traditional mining algorithms is not possible. So, to produce the large amount of data, the technologies of data mining need to be reconstructed. If not the technology that is present can only be applied to small scale IOT system which produce only small amount of data. To develop a data mining of KDD for IOT which produces high performance data, we use three key considerations that solves the KDD technology problems:

- Objectives (O): In this, first we need to specify the assumptions, limitations and measurements to solve the problem.

- Data (D): Data mining is concerned with the characteristics of data, such as size, distribution, and the presentation.
- Mining algorithm (A): The objective and data that is needed for data mining is specified above.

OPEN ISSUES OF IOT

The camera-based is used to detection and recognition technologies, which makes to concern of users, will make people uncomfortable. In recent year it is concerns of using data mining technologies to the data from IOT Fig 2.2 By taking example, now a day's almost all the companies are easily collecting the various data of consumer in the different sources or devices and by using this data mining technologies to get data that is helpful to marketing tactics to increase the volume and sales, many consumer don't like to have their privacy, such as shopping behavior, collected. The other example is the application to the health care in smart home or hospital. By this the sensitive data, that is the behavior of patients, should be protected. Some technologies have been proposed in recent year, such as randomly delay [6] and encryption.

Sound sensor is the alternative solution, RFID, or other wearing sensors. The new problem arises the new sensor or RFID technologies this can benefit the mining module of the IOT; there they will bring up the challenges. Data mining technologies can still recognize the identity of data is collected by the IOT. Data mining provides to add additional values for the collected data by IOT. Here we need to deal with, how to protect this information and how to avoid undesired influences while we are enjoying these benefits.

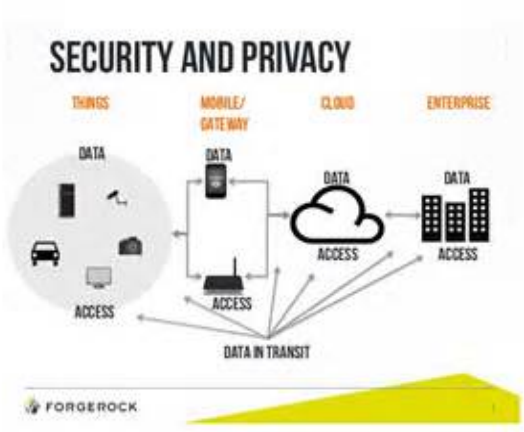

Fig.2.2 Privacy and Security in open issue of IOT

Open issues of IOT: the major 6 critical open issues are as follows:

1. How are devices named and organized?
2. How do devices communicate with each other?
3. How are devices tracked and monitored?
4. How is performance measured and optimized?
5. How is security and privacy safeguarded across billions of connected things?
6. How will these devices be maintained?

III.CONCLUSION

In this paper, a brief idea of applying data mining for IOT is given. Handling large scale data had become a big problem so we use many methods to overcome these kinds of problems. The open issues on privacy and security that we are facing in day to day life are presented.

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," Ad

Hoc Networks, vol. 10, no. 7, pp. 1497–1516, 2012.

[3] N. Ali and M. Abu-Elkheir, “Data management for the internet of things: Green directions,” in Proc. IEEE Globecom Workshops, 2012, pp. 386–390.

[4] R. G. Baraniuk, “More is less: Signal processing and the data deluge,” Science, vol. 331, no. 6018, pp. 717–719, 2011.

[5] V. Cantoni, L.Lombardi, and P.Lombardi, “Challenges for data mining in distributed sensor networks,” in Proc. International Conference on Pattern Recognition, vol. 1, 2006, pp. 1000–1007.

[6] C. Efthymiou and G. Kalogridis, “Smart grid privacy via anonymization of smart metering data,” in Proc. IEEE International Conference on Smart Grid Communications, 2010, pp. 238–243.