ISSN: 2393-8528

IJICSE
Open Access

Contents lists available at www.ijicse.in

International Journal of Innovative Computer Science & Engineering

Volume 5 Issue 1; January-February 2018; Page No. 01-05

A Survey on Uses of Data Mining Technology for IoT Services

Jitendra Joshi

JVW, University, Jaipur, Rajasthan, India.

ABSTRACT

Internet of things is a technological field that is limited to storage and computation in internet. It contains objects with sensors having softwares and network connectivity. It is really an impossible task to connect all the devices to internet, but it became possible through the use of internet of things technology. In near future IoT can change our lives very dramatically through its various applications in real life. The data gathered or generated through the applications of IoT is very useful in the daily life. Its really very valuable. This valuable and useful information can be more convenient through the use of data mining. Data mining can make IoT more and more efficient. It will make the system smart. In this paper we will discuss about the IoT, uses of IoT applications, the data mining techniques that can be used in it and the issues, challenges and potentials of using it.

Keywords: Internet of Things, data mining, Computation

INTRODUCTION

When the first time IoT has been introduced by the Kevin Ashton in 1999, this technology is considered as the technology to connect fundamental networks to networked objects. All the things or devices can be connected through IoT technology in internet. The devices are identified automatically. The whole idea of IoT is that the things can be identified automatically and they can communicate with each other, and they can even make decisions by themselves. Recent evolution in internet is based not only on collect information that is collected from the environment but also it uses the Internet standards that already exist to provide the services. IoT is a real world technology in which there are small devices that are called things. They have small storage and small battery or energy and limited capacity of processing.

There is a good progress in communication in last decade and the machine to machine communication and decision making is also increasing very dramatically. It was worth $44.0 billion in 2011 and is grown up to $290.0 billion in last 7 years. Thus, researchers from different industries, research groups, academics, and government departments have paid attention to revolutionizing the internet, by constructing a more convenient environment that is composed of various intelligent systems, such as smart home, intelligent transportation, global supply chain logistics, and health care. IoT can be viewed in different perspectives through different research that is carried out last few years. Different surveys are carried out on challenges in IoT, issues faced in implementation, applications of IoT and standards and smartness of IoT. These are all good surveys, those can be referenced in future surveys on this technology. So that enhancement can be made in future implementations of different applications.

Atzori and his colleagues has discussed about the IoT and its applications. There a recent study in which they have given a five-layered architecture to describe the design of IoT.

The five layers are: Perception Layer, Network Layer, Middle layer, application layer and business layer. Some of the studies described the infrastructures and things or objects of IoT. Some studies also discussed about the using of intelligence in the devices of IoT, they called them as “smart objects” (SO) and are assumed capable of being identified, sensing events, interacting with others, and making decisions by themselves. These devices or objects can communicate with the environment, they can collect the data by sensing. This data can be used in further decision making process without the intervention of the human. One of the most important questions

Corresponding author: Jitendra Joshi

that arise now is, how do we convert the data generated or captured by IoT into knowledge to provide a more convenient environment to people? This is where comes the knowledge discovery in databases (KDD) and data mining technologies, for these technologies provide possible solutions to find out the information hidden in the data of IoT, which can be used to enhance the performance of the system or to improve the quality of services this new environment can provide. A numerous researches are therefore focusing on using or developing effective data mining technologies for the IoT.

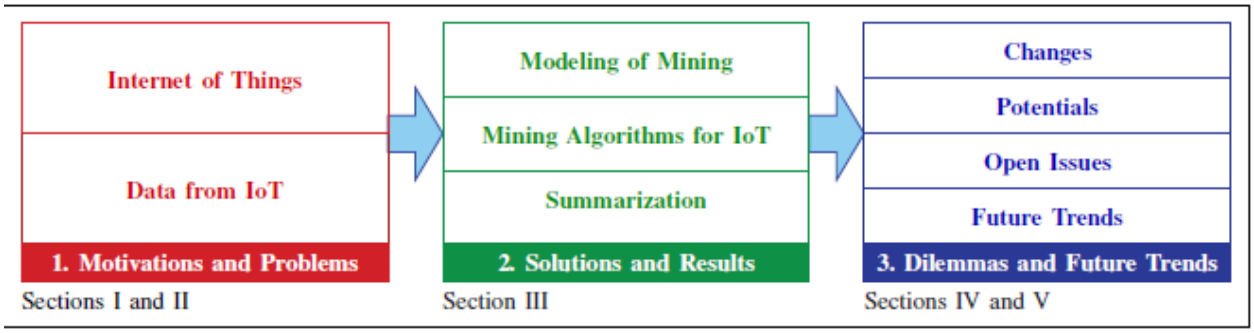


Figure 1: Levels of the Survey

All the things in IoT collects the data and that can be transformed into decision making by extracting information from the data using knowledge discovery. There are some technical issues and challenges on how to handle these data and how to dig out the useful information have emerged in recent years. These challenges and issues are handled when the data mining is used for KDD. This process is called as data about things. This is the data that describes things. Information gathered from this data can describe the location, identity and process and "data generated by things" to refer to data generated or captured by things. Normally, the former contains data that can be used to optimize the performance of the systems, infrastructures, and things of IoT whereas the latter contains data that are the results of interaction between humans, between human and systems, and between systems that can be used to enhance the services provided by IoT.

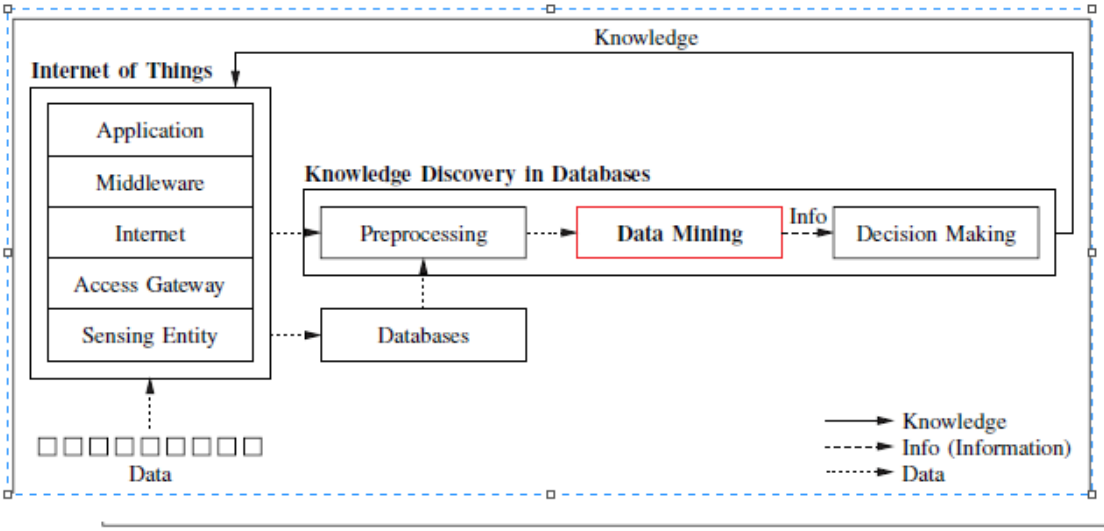


Figure 2: Knowledge Discovery in Databases

DATA MINING FOR IOT

This section describes the interconnection between some of the techniques used that are KDD, data mining and cloud computing. We have also given some basic example to show the interconnection between them by using a framework. After that, a detailed analysis and summarization of mining technologies for the IoT will be given.

Basic Idea of Using Data Mining for IoT:

Data collection is easier than the analysis of data. Data analysis is actually a smart process. It's a smart use of the data that is gathered from different sources. The explosion of data will certainly become a serious problem of IoT. Till now some of the studies have discussed about the efficient and enhanced analysis tools and techniques of analyzing the data collected. Without effective and efficient analysis tools, we, and all the systems, will definitely be submerged by this huge load of data. When KDD is applied to IoT, there are two perpectives that can be applied one is hardware and the other is software. Distributed database and cloud computing are the technologies that can be applied in hardware perpective of knowledge discovery. In the second, that is software we have to use the data mining tools and techniques. There are some data mining tools, that are designed to analyze the data and they can run on a single system. In the circumstances of big data, it is almost certain that most KDD systems available today and most traditional mining algorithms cannot be applied directly to process the large amount of data of IoT. To development of a high-performance data mining module of KDD for IoT, there are three key considerations in choosing the applicable mining technologies for the problem to be solved by the KDD technology objective, characteristics, and mining algorithm.

- Objective O: in the first consideration, we have to first specify the problem. In this there are assumptions, limitations, and measurements of the problem to define the problem to be solved. So that, the objective of the problem can be made clear.

- Data D: Another important concern of data mining is the characteristics of data, such as size, distribution, and representation. Different data usually need to be processed differently. Although data coming from different problems, say, *Di* and *Dj* , may be similar to each other, they may have to be analyzed differently if the meanings of the data are different.

- Mining algorithm A: there are two algorithms that are shown below. One is for data mining framework and the other is called k-means algorithm. As we have already shown the data and the objectives. Which means we have defined the problem and we have the data, that is to be analyzed, so now the data mining technique can easily be implemented. For instance, from the characteristics of data, if the amount of data exceeds the capability of a system and if there is no feasible solution to reduce the complexity of the data, then a novel mining algorithm is definitely needed; otherwise, the current mining algorithm suffices. Another consideration is related to the property and objective of the problem itself. If a novel mining algorithm can enhance the performance of a system, then the new mining algorithm is also needed. An example is the clustering algorithm for a wireless sensor network, which needs to take into account the load of computation, but most traditional clustering algorithms simply ignore this issue. "Now that the objective of the problem is decided, the characteristics of the input data are understood, and the particular goals of mining and the mining algorithms are chosen".

Algorithm 1 Unified Data Mining Framework

```
1  Input data D
2  Initialize candidate solutions r
3  While the termination criterion is not met
4      d = Scan(D)  [Optional]                    S
5      v = Construct(d, r, o)                     C
6      r = Update(v)                              U
7  End
8  Output rules r
```

Algorithm 2 k-means algorithm

```
1  Input data D
2  Randomly create a set of centroids c
3  While the termination criterion is not met
4      v = Assign(D, c)                           SC
5      c = Update(v)                              U
6  End
7  Output centroids c
```

Clustering for Services of IoT:

To make the system smart this is the important process that is developed. All the data gathered is stored in the clusters and then the system can make the decisions by themselves so that it can provide good services, such as detecting the falling event of older people by using a smart home system.

Classification for Outdoor Services of IoT:

Researches on using classification for the IoT can be divided into two categories: outdoor and indoor, depending on the region. One of the nightmares occurred frequently in our daily life is the traffic jam problem, especially when you are living in a big city. "More and more researches therefore are focusing on the traffic jam problem, by using mobile devices or smart phones to interchange information to avoid getting stuck on road, such as traffic forecasting services and

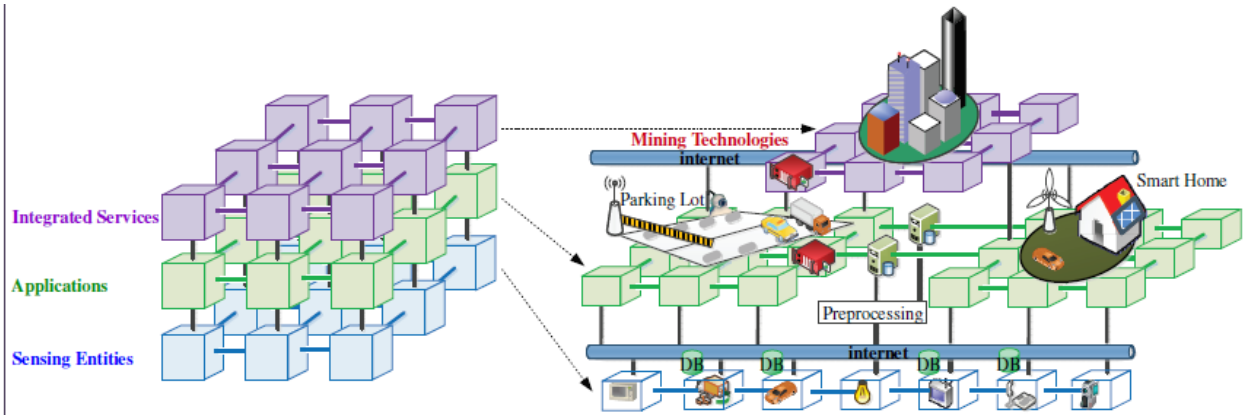
accident detection. Since getting information about the traffic situation via the internet and social network has become a trend, a driver guidance tool was developed by integrating the location information provided by the GPS, geographical information tracking from vehicles, and other information that can be collected from the internet to predict the future traffic situation". The decision tree classification algorithm is then used by this kind of system to predict and suggest the routing path for a driver, based on real-time information, historical data, and so on.

Which Mining Technologies Are Applicable:

Now that a unified framework has been presented to give a systematic description of data mining algorithms and how they are applied to IoT, the question that arises is, what mining algorithm is suitable for the development of a high-performance system or for the provisioning of a better service in different IoT environments that we may encounter. Fig. 6 gives a classification matrix to help us differentiate clustering, classification,

Changes Caused by IoT

The changes caused by IoT are discussed with different perspectives that are thing-oriented, internet-oriented, and semantics-oriented. "A very simple way to differentiate a device is if it can automatically become part of the IoT; that is, if it can be automatically connected to the internet, either directly or indirectly through other devices disregarding its capability. Another change is that devices are made smaller and smaller, implying that the lifetime of battery can be prolonged and thus more functions can be integrated into each device, such as integrating RFID and wireless sensor into a single device". These changes make it possible for everything to upload the data it collected to the server or application. We can then imagine that all the things on the planet being connected together will come true in the foreseeable future. Owing to these changes on things, data mining technologies are now able to enhance the performance of devices, to give them basic autonomous or cognitive ability to make decisions by themselves (e.g., to give air-conditioner the autonomous ability to automatically adjust the temperature of a room) and to interact with other things (i.e., machine to machine).



Open Issues of IoT

Although the focus of mining problems for the IoT differs from that of the traditional mining problems, they still inherit many of the open issues of the traditional mining algorithms and problems in the development of the mining algorithms for IoT, such as scalability for large-scale data set. There is no need to say that many other open issues need to be addressed in the design of the mining algorithms for IoT.

CONCLUSIONS

In this paper, the discussion is on applying data mining techniques and tools to enhance the functionalities of IoT. The data mining techniques include clustering, classification, and frequent patterns mining, from the perspective of infrastructures and services. The analysis and discussions of the implementation, challenges and issues of data mining technologies is also given in the paper. To make understand the discussion, we have also discussed the basics of IoT. The basics include infrastructure, architecture and applications. Objectives and characteristics of the fully integrated system are also given.

REFERENCES

1. K. Ashton, "That 'Internet of Things' Thing," 2009, RFID Journal
2. Auto-ID Labs, Massachusetts Institute of Technology, 2012,

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

4. Doukas, Charalampos, and Ilias Maglogiannis. "Bringing IoT and cloud computing towards pervasive healthcare." Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on. IEEE, 2012.

5. 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.

6. M&M Research Group, "Internet of Things (IoT) & M2M communication market - advanced technologies, future cities & adoption trends, roadmaps & worldwide forecasts 2012-2017," Electronics.ca Publications, Tech. Rep., 2012.

7. D. Bandyopadhyay and J. Sen, "Internet of things: Applications and challenges in technology and standardization," Wireless Personal Communications, vol. 58, no. 1, pp. 49–69, 2011