OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

DEVICE DISCOVERY AND CLASSIFICATION FOR INTERNET OF THINGS

Mrs.I.Priyadarshini[1], Mrs.A.V.Taware[2]
Assistant Professor, *Department of Computer Engineering, K.K wagh Institute of Engineering Education And Research(SPPU Pune), Nashik,India*
Assistant Professor, *Department of Computer Engineering, K.K wagh Institute of Engineering Education And Research(SPPU Pune), Nashik,India*
ipriyadarshini@kkwagh.edu.in [1] **,** *avtaware@kkwagh.edu.in*[2]

Abstract: Internet of Things allows everything around the world to connect to Internet .These include devices such as thermostats, utility meters, headset with blue tooth enabled, irrigation pumps, sensors or control circuits for electric car engine. It connects things embedded with sensors like devices, appliances and machines to the Internet and allows these things to communicate and exchange data. The number of connected device on IOT will be huge .By the year 2020 it is predicted that the number of devices connected to the Internet will be around 50 billion. There exist a vast diversity of the things being connected to Internet .Each device has its unique rules and standards for communication. It would become difficult for any application to integrate a new device as well as in case of any security threat in network it would be difficult to identify the device which causes the suspicious traffic, In order to solve the above problem there exist a need to automatically discover, detect and perform classification of device that are communicating inside the network.

Keywords: Internet *of Things, Device, Security, Classification.*

---∴∴∴---

I INTRODUCTION

Internet of Things has a wide range of application. It is used in various domains like home, city, health, industry, environment, energy management, agriculture, etc. It is used in smart home application for smart lighting that helps in energy saving by switching lights on/off according to occupancy in room and it can remotely control any devices or appliance .Incase of cities it can be used for smart parking to provide status on available slots. It can also provide smart health monitoring system. In Industry it can be used to detect and diagnose any fault in machine. It can serve the environment for forest fire detection, weather monitoring, air and noise pollution. For retail it can be used for smart payment, inventory management etc. For agriculture it can be used for smart irrigation. These were just some of the application of IOT and there is many more usage of IOT.

The Things that is, the device used in IOT are of varied type and each has a unique behavior and communication standard. There is a need to classify the device to facilitate integration of an application with a device easier and also to protect the device from security vulnerabilities.

This paper is organized as follows. Section II explains about the things in IOT its general block diagram along with different types of IOT devices. Section III explains about various communication protocols used by the IOT devices and Section IV explains about various methods used for device detection and classification and Section V briefs about the use of data mining for device recognition and section VI contains conclusion.

II THINGS IN IOT

Things in IOT refers to various device in IOT that is capable of sensing ,monitoring, controlling and exchange data with other connected device or send data across the network. Figure 1 shows the generic block diagram of an IOT device. An IOT device has several interfaces to connect to other devices of both wired and wireless. It has input output interface for sensors and actuators. Memory and storage interface, Audio/video interface and processor. IOT devices are of varied type which uses different communication protocols. All these device generate a lot of data which could

be processed and useful information can be obtained .Figure 2 shows some of the IOT device types. All these devices are heterogeneous in nature.

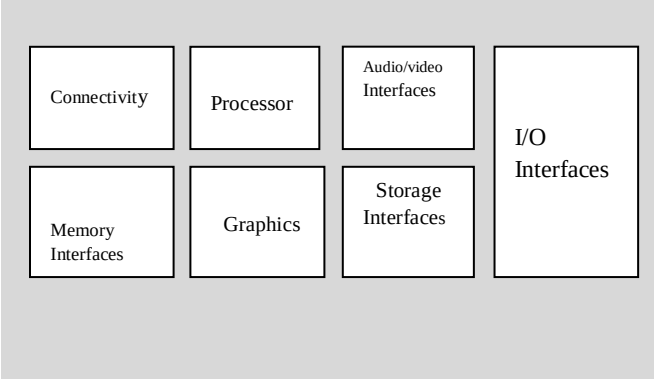

Figure 1 Generic Block Diagram of IOT Device.

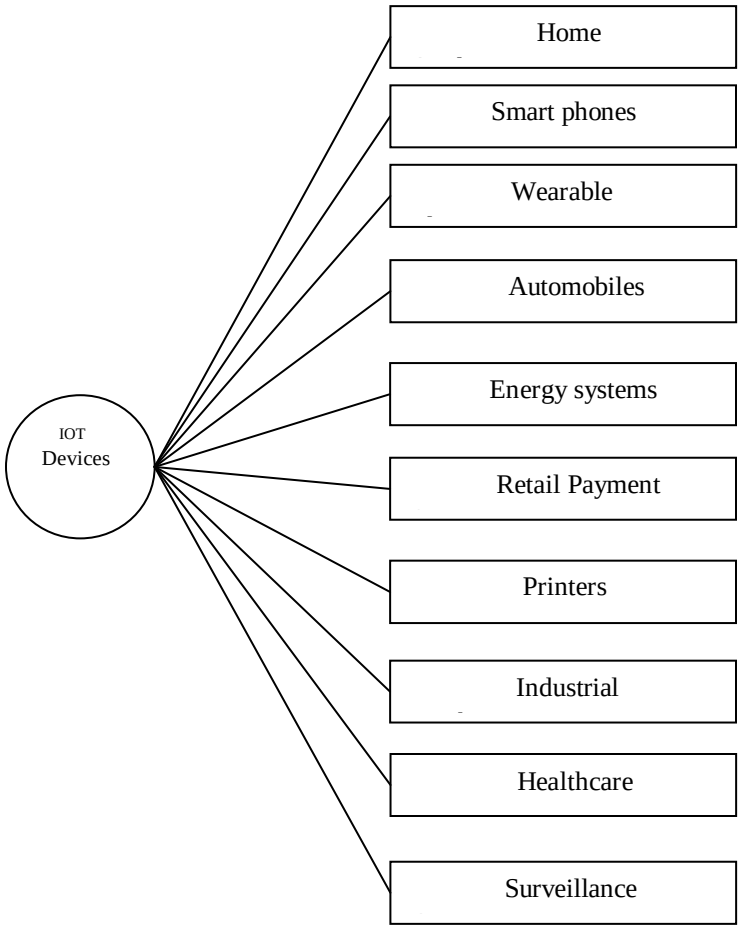

Figure 2 IOT Device Types

III COMMUNICATION TECHNOLOGY

There exist many communication protocols for IOT. Each device in IOT uses communication protocols to talk to one another. Some of the communication protocols like Wi-Fi, Bluetooth, Z-Wave, Zigbee, NFC its frequency band, range and its usage are listed in table.

TABLE I COMMUNICATION PROTOCOLS

Communication Protocol	Frequency Band	Range	Usage
Wifi	2.4GHZ to 5 GHZ	100 m	School, library ,phone, laptop
Bluetooth	2.44 GHZ	0 -10 m	Wireless Headphone, smart watches, wireless Keyboard
Z- wave	1 GHZ	30 m	Fire alarm, Fan ,remote control front door lock, thermostat
Zigbee	2.4 GHZ	100 m	Sensors ,industry
NFC	13.56 M Hz	4 CM to 10 CM	Car wallet, ID Badge, Bike

IV DEVICE DISCOVERY

As more and more devices are connected to the network it creates a threat to security .Once connected devices are properly discovered and profiled, anomalous activity detection can be used to continuously analyze device behavior and compare it to what should be expected for the particular device. Then appropriate protection measures can be taken.

A. Device Data

All devices inside a network send data to each other .These data is sent in the form of XML file or JSON format. These data file can be captured and after preprocessing and feature selection a device dataset can be prepared for training the system.

There are many algorithms used for device discovery in IOT .This section will brief about the various algorithm used for device discovery.

A. TF-IDF

T.f-idf stands for *term frequency-inverse document frequency*, and the tf-idf weight is a weight often used in information retrieval and text mining. It is used in search engine as central tool in scoring and ranking a document's relevance given a user query. tf–idf can be successfully used for stop-words filtering in various subject fields, including text summarization and classification. Given a particular communication file for a device this algorithm can generate a list of devices and categories that have highest Tf-idf value for that device.

B. Levenshtein Algorithm

This algorithm uses a string metric for measuring the distance between two metric [1]. The Levenshtein distance between two words is the minimum number of single-

character edits (insertions, deletions or substitutions) required to change one word into the other. This method can also be used to find the device in database whose property value has a similar value of our new device.

C. synonyms Match

Similar device has similar property. This approach takes a property of a test device and search in the database for a device whose property has a same meaning to the test device's property. For example the property hue and color means the same. Using this approach a test device can be categorized.

V DATA MINING IN IOT

In order to perform an appropriate device discovery data mining algorithms and techniques could also be used for IOT [3]. All that is needed is a supervised approach for device classification. Data set is essential for performing training and to classify any new object according to its feature.

A. Decision Tree Classification

Decision tree is one of data classification technique used in data mining for supervised learning. In decision tree Each internal node represents a test on an attribute, each branch represents the outcome of the test and each leaf node holds a class label. A typical decision tree can be constructed using ID3 algorithm[5][2] where each internal node holds the property of devices like mobility, bandwidth requirement, battery requirement, memory, Wi-Fi enabled, blue tooth enabled e.tc. and the leaf node contains the device category like sensor, light, controllers, display etc.

B. K means

K means algorithm requires a number of clusters I.e. the number of categories of devices to be formed and it outputs a set of clusters. To calculate distance between data points and center, the nominal attribute are given a numerical value .for example If the device is Wi-Fi enabled it is given 0 else 1.Tthe following steps are performed: [4]

1. Choose k objects from dataset as initial cluster centered
2. Repeat
3. Re(assign) each object to the cluster to which the object is most similar based on the mean value of the object in the cluster
4. Update cluster means, that is, calculate the mean value of the objects for each cluster
5. until no change

The Performance of each algorithm can be measured by calculating the precision, recall and F-measure.

VI CONCLUSION

Internet of Things includes vast diversity of devices with different features. It becomes essential to devise a method which would automatically detect and categorize the device used in the network to facilitate easy integration of the device to any application and it also becomes essential to know the behavior of each device in the network to protect the device against any security threat. This paper has focused on different approaches that could be used for identifying a particular class of the device. These approaches may help the IOT developer make the connectivity easier and safer.

REFERENCES

[1] M Pedro R.J, Pejo. Luis Nunes, "Automatic Discovery and Classification of IoT Devices", in Information Systems and Technologies (CISTI),IEEE 978-9-8998-4347-9, 2017

[2] A. Joan Nolla Suarez, Ante Salcedo "ID3 and K-Means based methodology for Internet of Things device classification , International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE) IEEE, 2017.

[3] EMC, "Advanced Analytical Theory and Methods: Clustering," in Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, 1st ed.Indianapolis: Wiley, 2015, pp. 118-130.

[4] J. Twining, "Behind the numbers: growth in the Internet of Things," Platform with information from Cisco IBSG. March, 2015

[5]Chun-WeiTsai; Chin-FengLai;Ming-ChaoChiang; LaurenceT.Yang, "Data Mining for Internet of Thin A Survey IEEE Communications Surveys & Tutorials, Volume: 16, Issue: 1Pages: 77 – 97, 2014

[6] D. Evans, "The Internet of Things:How the Next Evolution of the Internet Is Changing Everything," Cisco IBSG. April, 2011.

[7]E. Fleisch, "What is the internet of things? An economic perspective,"Economics, Management, and Financial Markets, no. 2, pp. 125–157,2010

[8]J. Ramos, "Using tf-idf to determine word relevance in documentqueries," in Proceedings of the first instructional conference on machine learning, 2003.

[9] J.K. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, no. 1.Netherlands, 1986, pp. 87-92

[10] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, 1967.

BIOGRAPHY

Mrs.I.Priyadarshini holds M.E(Computer Engineering)from K.K.wagh Institute of Engg.Education and research,SPPU university Pune.She is Working as an Assistant professor in Department of computer Engineering in K.K.wagh Institute of Engg.Education and research.Nasik.

Mrs.A.V.Taware holds M.Tech(Information Technology) from,RGPV university Bhopal.She is Working as an Assistant professor in Department of computer Engineering in K.K.wagh Institute of Engg.Education and research.Nasik.