International Journal of Advanced Science and Technology
Vol. 29, No. 3, (2020), pp. 3223- 3237

Internet of Things based e-Healthcare System: An Useful Review on Critical Issues and Challenges

A. Mohan Krishna
Research Scholar,
Department of Computer Science and Engineering, Lingaya's Vidyapeeth
Faridabad, Haryana, India.
amkrishna@hotmail.com

Amit Kumar Tyagi
School of Computing Science and Engineering,
Vellore Institute of technology, Chennai Campus
Chennai, Tamilnadu, India.
amitkrtyagi025@gmail.com

Abstract

Due to recent development in smart/electronic devices or computing devices, a lot of data (called Big Data) is being generated. Such data in simple terms is called as "Large Data" or "Big Data". Today's Big Data is highly useful for predicting future (with respect to an application) after having a proper and valid analysis. Together this, several researchers have tried to make efficient analytics tools (or techniques like data mining or machine learning) to analysis this Big Data, but some tools failed due to the volume and veracity feature of the big Data. So, there are several problems, issues and challenges likely to be rectified in the near future. This article gives an insight to where and how big data can be used in health care and other possible systems/ sectors in near future. This work also provides an understanding by enumerating some challenges big data analytics has in healthcare sector.

Keywords- *Big Data Analytics, Challenges in Big Data, Smart Devices, Data Mining and Machine Learning Techniques.*

I. INTRODUCTION ABOUT DATA

Data is set of values and statistics, which can be used for analysis-purpose. The word 'data' originated from a Latin word known as "datum". Different types of data are included in [17]. Data can be collected from primary sources like text book or through observations, surveys, or by doing experiments. Data may be raw, unorganized facts that can be converted into meaningful information with the help of data pre-processing. Note that data pre-processing include cleansing, transformation and integration operations. Several countries have recognized the importance of data and particularly e-healthcare applications that may costs/account for a good percentage of its economy (for a nation). The healthcare industry historically has generated large amount of data, driven by record keeping, compliance & regulatory requirements and patient care. While most data is stored in hard copy form, the current trend is toward rapid digitization of these large amounts of data[3]. Further, data has helped many governments in implementing several social welfare schemes to their citizens. According to report [4], one third of the generated data in e-healthcare applications today is completely wasted and is not supplying anything to improve healthcare outcomes. This wastage in healthcare include worthless assistance, wastage in administrative side like huge prices, fruitless and replicated documentation, weak services, scam and missed prevention chances etc. Big data are primarily the data sets with huge amounts of data which can only be dealt with advanced techniques and powerful machines, else it would have taken several years to process. Communication without any distortion is a major requirement in healthcare sector. This increases

ISSN: 2005-4238 IJAST

participation of doctor and patients in the process of health care. All this can be achieved using Big Data techniques as a solution. In addition, we can also detect security threats with the implementation of Big Data which always aims to extract value from data economically. Although the technique provides numerous benefits, resistance to evolve into Information and Communications Technology (ICT) based care from traditional mode, the longitudinal impressions of health information, unavailability of standards and security challenges together make it unviable to make use big data in this sector. The different initiatives which used the advantage of big data in health care sector are listed as below:

a) Asthma Polis: It is a GPS enabled inhaler monitors inhaler usage of an asthma patient. Whenever a patient uses the inhaler, the apparatus collect the details about geo-location and time of usage (the place and time the inhaler was used) and reports such data to a web application. Thus the usage details of an asthma patient are made available for both patient and also to the doctor. Patient can utilise such information to take preventive measures whereas such details would help the doctors to decide personalized treatment.
b) Battling the Flu: The centre disease control collected large amount of flu records. These reports include the information like disease description, previous therapy adopted, and the status of this medication etc. This information helps the patients to take precaution and also help the caregivers with information such as what vaccine to use for the nausea and which medical assistance can help the patient in their rehabilitation etc.
c) Flu near You: This is an application which gathers the symptoms of the users before they fall ill in reality. Thus, by collecting data before people falling sick, an activity map is generated which would help the people to take preventive measures from disease.
d) Help, I have the flu: This is an application that fetch the data from social media and the person responsible for carrying the flu.
e) Go Cap: Go Cap is an insulin pen invented by a company named Common Sensing. It can capture the daily insulin usage amount, and also the specific times when it has been used. This data help healthcare experts to identify problems before they become serious and also help in adjustment of dosages if needed.
f) The Moon Shots Program: The main intention of this program is to enhance survival for different cancers. Collaborative endeavour among different companies and cancer analysers are involved in this program.

As discussed above, we have reviewed several issues/ wastages in analysing or collecting Big Data. By collaborating the public and private sectors (industries and organisations), it is possible to enhance the performance in healthcare system. Predictive models can be constructed by combining data from different sources which help to reduce the overall cost and enhance quality of therapy by using evidence-based and preventive-based medicine. We are attempting to define the scope of the Big Data in near future. Big data in healthcare is essential because its volume, diversity in data type and speed at which it can be managed. Note that data selected for e-healthcare applications include clinical data(prescription, insurance, pharmacy), patient data, machine generated/sensor data, less treatment oriented data(emergency care data, articles in journals). Some advantages in e-health care (of big data/ genomic analytics) are:

- Detecting diseases at earlier stage.
- Prevention of disease.
- Detecting health care fraud.
- Predicting: patients who prefer surgery, who may not get profit from surgery, chances of advancement in diseases etc.
- Evidence based care
- Reduced cost of healthcare
- Increases participation of patient in the process of recovery.

ISSN: 2005-4238 IJAST

- Improves public health surveillance
- Reduces mortality rate
- Decreases the communication gap between healthcare providers and patients.
- Early detection of security threats and fraud in healthcare
- Improves the Quality of Care

The above data is collected from different applications like transport, financial, geographical, scientific, etc. Data Masking replaces sensitive data elements with an unidentifiable value. Electronic Healthcare Predictive Analytics (e-HPA) have the potential of making decisions to predict and enhance the health of a patient undergoing treatment. Moreover, unique features of big data such as heterogeneity, incompleteness, data privacy, ownership, timeliness and longevity specifically in health care make the subject more challenging. In this work, we are exploring to find the answers to following basic questions:

- From where health big data originates?
- How can welfare arise out of reality mining techniques?
- Implications and applications of Big Data analytics in healthcare systems?
- Could the implementation of new tools such as HER?

Rest of the paper is organized into seven sections. Facts related to evaluation to data, Big Data revolution are explained in Section II. Importance of data and big data are covered in detail in Section III. Then, we have discussed several existing tools to refine this data for future (further) use in section IV. Then, we discussed "why these existing tools are not enough to refine this data" in section V. Further, several critical issues like security and privacy in medical devices (internet of things) are discussed in the same section. Internet of things based healthcare with respect to Big Data analytics have several critical issues to address and the obstacles in implementing data science are described in Section VI. Finally, we have concluded this work by summarising the review of e-healthcare systems and listing out the scope for future enhancements in section VII.

II. EVOLUTION OF DATA

Data is growing day by day in terabytes from millions of devices (connected in different – different applications). For several years, data processing functions used to be performed manually on paper. Electronic Discrete Variable Automatic Computer (EDVAC) is considered to be the first computer to use magnetic tape for storage developed around nineteen-fifties and then main frame computers also came into existence. Third generation languages such as COBOL, ALGOL, BASIC were mainly stored data in file systems. Database Management Systems emerged in the 1960s to address the problem of storing and retrieving large amounts of data. Different varieties of data like audio, video, images, text also came into existence [15]. Initially data rate was increasing at a slow pace, then in the twentieth century data started increasing exponentially, and later in 2015 the growth of data became unexpected rate.

Small organizations started maintaining their data in the personal computer having Windows operating System with some kind of databases. Office suite of Microsoft was released with MS-Access database and the installations of databases such as Oracle, paradox, Sybase SQL Server increased exponentially. In 1990's, Pentium processor, Internet explorer, Hotmail, Java language, JavaScript language were rolled out. Smart phones were also released during that period. Advent of Internet, rapid penetration of mobile phones, acceptance of Information Technology in non-computational subjects such as Biology, Medicine, Finance, Mass Communication have increased the volume of data.

ISSN: 2005-4238 IJAST

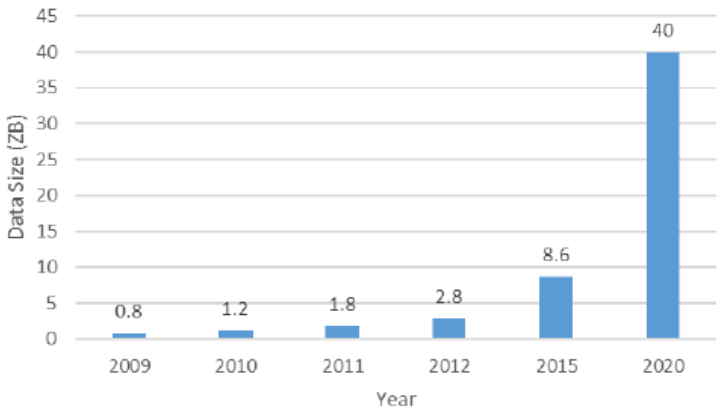

Figure 1: Data Generated in Health Care Applications (2009-2020) [5]

Client server installations were slowly replaced with web based applications due to user friendliness provided by the browsers such as Chrome, Internet Explorer, Firefox etc. According to Liang Hong et al [5], Big Data in healthcare sector exceeded 150 exabytes after 2011, and it is also stated that data size in healthcare sector estimated to be around 40ZB in 2020. Further, it may be noticed from Figure-1 that, data size in 2009 was just 0.8 ZB and increased about 50 times by 2020. This Big Data contains following dimension (V's), such as:

- Volume: It represents size of the data
- Variety: It represents number of types of data
- Velocity: It represents agility of data processing
- Veracity: It means data assurance, i.e., outcomes of prediction are error-free and honest.
- Value: It means the value of the data.

Data may be available multiple forms like structural and semi structural including instrument reading health care or medical records. Unstructured include office medical record, doctor prescription, MRI, other images, admission, discharge report._Note that initially the scale of data rate was in Mega Byte, Giga Byte, Tera Byte, which was able to store and manage by traditional database management system.

But later the data was increasing in the scale of Petta Byte, ExaByte, YottaByte which traditional database systems could not handle effectively. Mainly three V's namely volume, variety and velocity differentiated data and Big Data. Volume referred to the massive amount of data generated from multiple sources which made Scientists to explore new avenues for storage and processing models. Variety referred to heterogeneous and incompatible data formats received from different sources in variety of forms. Velocity refers to the speed at which data is generating and accessing. Figure 2 characterizes big data in healthcare into 5V's includes volume, velocity, variety, veracity and value. More V's or more features of V's can be found in [18]. Precisely, Big Data indicates huge, complex datasets which cannot be handled by traditional data management systems (in terms of storage, management and secure-timely processing). The application of big data includes finance, marketing, health care etc.

ISSN: 2005-4238 IJAST

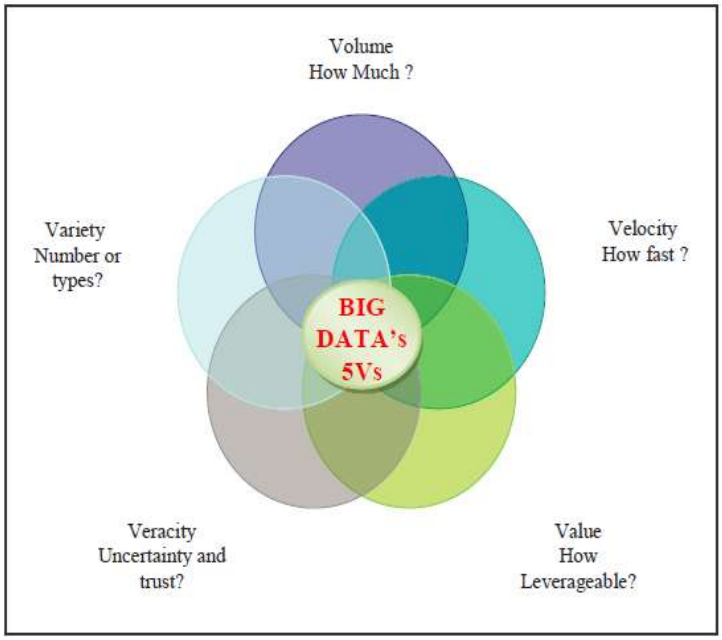

Figure 2 Big Data's mainly Five 'V' for Healthcare [7].

A. Big Data: Revolution

A large amount of data (In 2005 Roger Mougalas from O'Reilly coined this term Big Data for the first time [15, 16]) produced by IoT devices, is a challenge for all software testing Services Provider (SP) to provide enough security to the IoT devices. The increasing cost, shift in clinical landscape, shifts in provider reimbursement trends are the main issues which enhance the demand of big data. The major sources of big data in health care sector are;

a) Clinical data like electronic medical records, MRI report, prescription, etc.
b) Claim and cost data like utilization of care, cost estimate.
c) Pharmaceutical Research and Development (R&D) data like clinical trials, screening libraries.
d) Patient behaviour and sentiment data like patients preferences, purchase history etc.

Many firms, industries, organizations and institutions provide data-driven informatics derived from integrated data sets. Improvement in technology helps to overcome the disadvantages of traditional system in storing, compiling and sharing information securely. Third Party tools available for removing names and other personal information from healthcare records and government also provides good support for big data revolution. For example, government authorities reevaluate the cost of the drugs based on the reports generated by the medical agencies that are gathering and examining clinical data on expensive new drugs. These initiatives definitely help to promote the data release and accessibility, provide data standardization and improve the quality of the big data system. The new value path ways of big data in healthcare are

i. Right living: Encouraging the patients to follow good diet, practicing exercises to stay healthy, initiating proactive role in their own treatment, and to prevent from disease attacks.
ii. Right care: Making sure that patient gets timely, appropriate treatment in affordable way. Right care is a team activity where all caregivers should have same information and every one need to work toward meeting the final objective and avoiding suboptimal paths.
iii. Right provider: Which assure that patient always treated by highly skilled professionals?

ISSN: 2005-4238 IJAST

iv. Right value: Include different ways by which the healthcare value can be improved by reducing the cost and enhancing the quality.
v. Right innovation: This includes invention of new treatment and therapies.

Note that since 1975, a different type of data 'genomic data' comes under scope and also intend for preserving people personal information. But in previous decades, sufficient and efficient tools were not available to refine this critical or healthcare data. The traditional medical management system and fee-for service payment structure must replace with new methods to improve the use of big data in health care. Several measures are initiated in health care sector for using the big data and some of the examples are

a) Asthma polis which monitors inhaler usage by asthmatics.
b) Ginger.io where the patients are monitored with the help of mobile phones and helped with behavioral treatments.

B. Overview of Big Data in Healthcare

In a raw description, big data is nothing but a huge amount of complex data set that cannot be managed in affordable time with conventional techniques and computational power. The big data analytics is nothing but statistics in mathematical terms. From a health care view point, Big Data is a collection of techniques, technologies which undergo procedures to generate expected, process existing, analyze huge amounts of data based upon previously retrieved information from patients. The sources of big data in healthcare include the following:

i. Machine Created Data: Data retrieved from remote sensors, smart meters, wearable devices etc come under machine created data. Machines used in healthcare systems give these as output signals.
ii. Biometric Data: It is the information that is extracted from personal physical characteristics such as blood pressure, finger prints, retinal scans, signature, genetics, heart rate, pulse and x-ray and other medical images.
iii. Human Created Data: These are the information extrapolated by humans in the healthcare system. Data such as unbuilt and half built clinical data like case notes, hospital admission records, laboratory results, discharge summaries, electronic mails are noteworthy examples of this kind.
iv. Transactional Data: These include data from finance and healthcare claims.
v. Behavioural Data: This is the information drawn from social media and other related surveys.
vi. Epidemiological Data: Statistical data, health surveys and disease registries are some examples of this kind.
vii. Publication Data: Data concluded after several research and medical references.

Protected Health Information(PHI) in their 2016 Breach Report, informed about first ransomware attack happened in U.S. hospital and also stated that several major healthcare organizations were the target of the hackers. Additional findings of this report were:

- 325 large breaches of PHI, compromising 16,612,985 individual patient records.
- 3,620,000 breached patient records in the year's single largest incident.
- 40% of large breach incidents involved unauthorized access/disclosure.

Some of the technologies to avoid the mentioned disadvantage of open medical data include

a) Restricting the access to open medical data.
b) Data de-identification technologies
c) Cryptographic method

ISSN: 2005-4238 IJAST

Hence, this section discusses evolution and revolution of data including Big Data in detail. Now, next section will be critically examining the importance of data and specifically about Big Data.

III. IMPORTANCE OF DATA OR BIG DATA

Generated data can be used to predict something useful either by doctor to cure a disease or by an organization like Planning Commission for implementing their policies. For example, if a person wanted to do transaction with the Government in nineties, the concerned Office Assistant used to check all the records manually to verify whether the user was authenticated or not. This was a very time taking and also a hectic process. But, now days if the concerned Office Assistant types the account number of the person, entire details of that particular person are being displayed since Government has linked Aadhaar with several other applications such as PAN, Driving Licence, Bank Accounts etc. The new technologies (i.e., Aadhaar) made the human life easier and have been saving millions of rupees through direct benefit transfer schemes. Note that these linking of Aadhaar with other schemes is also generating a number of other large data-sets. Take another example from a private sector like an individual searching for a particular product and may not have purchased that product from that specific website. After words, if he visit any new web page, then the details of the product that he had searched earlier are getting displayed to attract the user to buy the product from their website. This is done with the help of prediction, or recommendation tool. If we talk in simple terms, data is required to make accurate predictions which make the human life easier or make human' being life safer and long.

As discussed above, we get to know Big Data's importance in several applications including healthcare and government. Data in health care can be categorised as the data which is directly related to patient and the data which are indirectly related to patient. The digitization of such data help to improve the quality, reduce the cost. Similarly, the government generates huge amounts of data every day. Some techniques are required for analyzing large data sets. It helps the government to provide value-added services to its citizens. Big data analytics can help in finding valuable decisions by understanding data models and their relationship with the help of machine learning algorithms. The role of big data in health care is useful for patient oriented diagnosis, to identify the cause of spreading diseases in the initial stage, to monitor the quality of health centre, to enhance the quality of treatment methods etc. The health care system has a large volume of unstructured data, so it is impossible to do research and diagnoses without an appropriate tool or technique[9]. The role of big data in government includes addressing basic need quickly, providing quality education, to reduce unemployment rate. In healthcare architecture, electronic health record is a heterogeneous data set that is given as input to HDFS through flume and sqoop. Examination of this data is done with the help of Map-Reduce and HIVE by implementing machine-learning algorithms that helps to identify similar data patterns. This also helps in predicting the risk of patient health condition at the earlier stages and with the help of Hadoop the goal of effective citizen care management can be achieved. It is a complex system that constitutes of components and technologies to handle large scale data processing and analytics on it. Big data ecosystem consist of collecting the data from different sources, preserve this collected data in HDFS, process them with the help of Hadoop components like Map-Reduce, examine them using PIG and generate Business Intelligence reports such as patient scorecards. A big data life cycle contains data collection, data cleaning, data classification, and modelling and data delivery. HDFS (in big data analytics) this file system mainly include two nodes.name node which act as a master node and data nodes which refer to slave nodes. In summary, impact of big data analytics on the healthcare industry are:

- The perspective of the research institution and the hospital
- The perspective of the government or the public
- The perspective of patients and their relatives

ISSN: 2005-4238 IJAST

Hence as discussed above, we will see the importance of data in health care field. People are using rubber band to check the heart beat rate, machine to machine devices are used in most of the hospitals, and most of the sensitive operations are done with the help of machines. This process saves user's time. In a nutshell, this section discussed in detail the importance of big data including data in several applications. In the next section, we will be discussing about existing tools for refining the data or Big Data.

IV. EXISTING TOOLS TO REFINE THE DATA

As discussed above, Big Data can be defined as "huge volume, high velocity, and large variety information source that require new forms of processing to help in strong decision making, and process optimization" [1, 16]. Integration of all these details helps to create a complete view of every patient. By integrating various social media details, environmental information, and behavioural information related to patients, organizations are able to discover new correlations. The digitization of healthcare information helps to improve the quality of service, health care results, and also help in cost reduction. Hadoop is one such tool that is designed to process huge volumes of data and capable to integrate with map-reduce program. This framework provides an interface for the distribution of sub-tasks to several mappers and finally the gathering of outputs from the reducer. When tasks are executed, MapReduce tracks the processing of each server/node. MapReduce has the packages like Hbase, Hcatalog, Oozie, Mahout, Sqoop, ApacheHive, and Apache Pig . These packages provide unique strength to the system by performing analytics directly on data. The main applications of big data in health care include personalized treatment, utilization review, fraud detection etc. The prediction model include various steps like collecting data from various sources, identifying the prediction attributes, applying algorithm for these attributes and then use contextualization to retrieve the desired performance.

A. Drawbacks of Healthcare sector implemented by Data Science

The main challenge in implementing Big Data in healthcare sector is that the source data is improper and insufficient for making decisions, planning and strategically care process although it is beneficial in other contexts. Therefore, the data must be evaluated and integrated to extrapolate supplemental benefits using it. Hence, let us go through the obstacles of implementing data science in health care.

a. Resistance *to Change:* The healthcare process system ceases in adopting technology unlike other sectors like the banking, auto-mobile. Improper administrative support for Information Technology (IT), legal issues, the unavailability of required and proper skills for the implementation of the Information and Communications Technology (ICT) tools. Hence, the paper bound techniques are still thriving in the healthcare system. The paper bound system nevertheless does not support the augmentation of data generating from different sources in the healthcare system.
b. Fragmentation of Healthcare Data: The health care system is lagging behind the banking sector and other sectors such as the oil industry. This is due to information technology and allied learning changes, trust factor(since it is a sensitive issue), legal issues, the unavailability of required and proper skills for the implementation of the Information and Communications Technology (ICT) tools in the healthcare sector. Hence, the paper based mechanisms are still prospering in the healthcare systems. The paper bound system nevertheless does not benefit from the merits derived from the Information Technology.
c. Ethical Challenges: Ethical challenges such as data privacy, confidentiality, control over access to patient information, ownership of patient information, and commercialization of

ISSN: 2005-4238 IJAST

administration information are some of the factors that hinder effective communication between patients and health care providers. As a result, integrating health care data from different sources becomes challenging. Therefore, access to a detailed and complete picture of the patient during timely care has became an issue.

d. Proliferation of Healthcare Standards: Standards that allow systems, equipment, technology, and platforms to work together are accepted by the health care organizations but do not meet implement such standards quickly. For example, case reports, medications, diseases, and test titles and codes change with hospitals are not uniform due to continuation of old formats instead of adopting the standard formats.
e. Security and Privacy Issues: A main obstacle in bringing huge sources of health data into data sciences is about security and privacy of the consumer. Health care data is vulnerable to security Trojan threats, such as intangible and unacceptable disclosure of patient data, unauthenticated usage and destruction of patient information. Therefore, health care service holders are uninterested to share health information via digital healthcare systems.

B. What We Want?

As discussed above, machines are connected together to reduce workload of the doctors in near future, and providing efficient services to patients. Also, these devices are generating a lot of data in communication (in motion/ dynamic and static mode). A medical cyber physical system is a networked system to work and make the entire communication system to be more secured from end to end. Communication should be encrypted with suitable cryptographic mechanisms in order to prevent the leaking of sensitive information from such devices. These devices (internet connected) and machines (which are communicating together) need to protect from external and internal attacks. Further, Hardware security also need to considered, because it became critical when unknown user access the medical devices physically. For example, most of the IoT devices have hardware vulnerabilities, because they makes communications using some protocols, addressing schemes and sensors, and most importantly they are connected together with internet, that may infect by malware. Also, efficient analytic tools need to be developed for analysing the large amount of medical data which is in unstructured form.

Various technologies are in use to ensure security and privacy of big healthcare data. Most widely used technologies are: Authentication, Encryption, Data Masking, Access Control, Monitoring and Auditing. Also, some useful tools for big data analytics are Google Big Query, Map Reduce, Hadoop, Jaql, and Microsoft Windows Azure. Note that more tools for analysing big data can be found in [17]. In Summary, this section discussed several existing tools for analysing big data, also shown light on why these existing tools are not sufficient and efficient to analysis this big data. In the next section, we will be discussing several security issues in internet connected devices in the medical sector.

V. SECURITY ISSUES IN MEDICAL DEVICES/ INTERNET CONNECTED DEVICES

As discussed above, internet of things (or devices which are connected together via internet) have an essential role in e-healthcare applications generally referred as medical applications. Some of the critical issues concerning these devices needs to be investigated and providing security in Internet of Things (IoTs) devices is a critical and an essential issue for winning trust among people and also the organisations. Data of Internet of Things (IoTs) could be industrial, enterprise, consumer or personal but it must be secure and confidential against theft and tampering. Security is also a concern where data is transmitted securely over the Internet because same IoT app is used for storing the patient's health and also for procuring some item through online store. Issues like privacy such as leaking of personal data of

ISSN: 2005-4238 IJAST

users, data sharing with others also raises the issue of trust in IoT ecosystem [11]. Some of the security issues in medical device are listed below.

- Security and Privacy*:* As discussed in section 2, Internet Connected Devices/ Things can communicate together, and may share data back to respective service providers (or master: who build these devices), and compile data for third parties such as researchers, health care providers, firms, organisations, or even other consumers. Hence, the storing and sharing of information with smart devices/ IoTs devices bring new challenges for regulators, enterprises and consumers. Securing this information and preventing this information from unwanted leakage (by attacker/ user) is yet to be addressed perfectly in these internet connected devices [12].
- Internet of Things Ecosystem*:* The IoT revolution is still in elementary stage and is evolving rapidly. 'Things' (e.g., everyday objects, environments, vehicles and clothing) may have more and more information associated with them, and are beginning to sense, communicate, and produce new information while becoming an integral part of the Internet. In near future, market of IoTs devices might grow substantially, with an introduction of new business models that are being implemented in several applications for providing efficient services. These will also stimulate innovation and growth in areas such as components, devices, wireless connectivity, system integration and decision support tools.
- Device/ Physical Security: Connection devices are main components of IoT, which collects data and interact with other devices including human beings. When these devices collect data, then they may be vulnerable to physical security issues. Unauthorised physical access to connected IoT devices can happen even with robust network and that might create a problem of catastrophic system failure. Hence, proper security measures must be initiated to prevent unauthorised physical access. This kind of attacks generally occurs when a malicious user/ attacker permitted to disassemble a device. Weaknesses may also occur when Universal Serial Bus (USB) ports or other external ports are used to access the device using features intended for configuration or maintenance. This could lead to easy unauthorised access to the device as well as the data. These vulnerabilities can be avoided by following the counter measures:
    (i) Frequently changing the medium of data storage and properly encrypting the collected/stored data at server side, i.e., using secure cryptographic mechanism.
    (ii) Ensuring USB ports/ other external ports to be accessed only by verified devices and preventing malicious access of the device.
    (iii) Ensuring device cannot be easily disassembled and provide limit use of data to administration/ authorised users.
- Network Security: Networks are always vulnerable to hacks from long decades (i.e. before advent of IoT). Several hackers tried to occupy network for their financial use. We have large volume of IoT devices, which makes us to have a robust and secure network. In the past decades, several scientists and researchers had analysed the security components and observed that vulnerability in network is due to weakest link in processes while transferring the data (in source to final destination). A network can be exploited by malicious users/ hackers via remotely, i.e., without access of IoT device physically [13]. One possible better solution for proper security is to use Virtual Private Networks (VPN's). It secures a network by encrypting the data traffic that flows through them [14]. Note that VPN's does not ensure absolute security to a system/ network and still susceptible to Man In The Middle (MITM) attacks.
- Data Security: Data in IoT can be classified into two categories; stored data (data at rest) and data in the process of transmission (data in flight). For maintaining data integrity, we need high-level encryption of both the data types. But, with a large variety of devices and varying hardware specifications, it is impossible to create a one size that fits all standard data encryption process. Data which is highly sensitive like usernames, bank account details, and passwords are required to encrypt with more factor authentication processes for ensuring safety and security.

ISSN: 2005-4238 IJAST

- Operating System Security: Operating Systems (OS) are prime target for any malicious user/ attacker/ hackers. If a hacker gains an access to the Operating System of an IoT cluster, then attacker can exploit or compromises into a system. Note that recovering from an OS security breaches is very costly affair and also require a lot of time to restore an OS to full efficiency. Some times, it may lead to data loss and backups of a system can minimise the overall cost, whereas, it is impossible to accurately detect the date of a hacking or compromising. Further, it is also impossible to know the exact point from where a system need to be restored. Hence, with the increasing size and complexity of IoT devices, a more robust IoT security is required to identify and neutralise IoT specific security breaches in near future.
- Server Security: Now a day's, IoTs are working as smart devices and interacting with cloud servers. Denial of Service (DoS) attack is one of dangerous attack that affect servers, i.e., malicious users use a large number of proxy devices to generate fake requests to the server. It makes a server ineffective and unable to attend requests of real users. Several steps has been taken to protect server security include limiting the number of open ports and exposed services. Here, security configurability needs to be a primary issue to solve, i.e., when an IoT system is being developed with allowing systems to be updated remotely with proper encryption and validation of update files.

With IoT security analytics, remote monitoring systems, automated patching procedures, we can secure IoT systems. However, security challenges in IoT systems might grow in near future as lot of advancements are expected in IoT. That might increase complexity in solving any issue/ to create new/ updating existing security countermeasures to protect IoT technologies. Note that IoT security issues can be available with different nature and occur at different levels. So, every organization/ firm/ company in the IoT sector must ensure the security, privacy and experience of users so that producers can really take advantage of the benefits of the Internet of Things. Implementation of data sciences in healthcare sector enables security and privacy of customer concerns due to security and privacy concealing real-time data analytics will uplift welfare in healthcare.

In this section, we have discussed several popular issues raised with internet of things and also tried to provide maximum counter measures including possible solutions to each of the issues. In the next section, we will be discussing several critical issues and challenges noticed and rectified in IoTs (or Internet-Connected Things) with respect to healthcare applications.

VI. CURRENT ISSUES/ CHALLENGES IN HEALTHCARE WITH RESPECT TO BIG DATA ANALYTICS

There are numerous situation oriented and technical drawbacks in the contemporary healthcare systems and specifically in the care process. Healthcare is required to ensure welfare of human beings. The current system cannot cope up with the par levels as it had a lot of defects and changes are required in the factors like patient satisfaction, finance issues, hospital productivity, hospital security, patient safety, information encryption, regular standards and power requirements. All these issues have to be taken care for better healthcare system to be established. Prescription drugs are noxious to humans and therefore, a reliable system to limit the quantity of ingested drug may have to be introduced. Main flaw in healthcare IoT is that hospitals in countryside are not installed with welfare checking system that can evolve along with emerging technologies like IoT. Facilities such as neural health services, health monitoring subsystems are yet to be established in several remote places. One another problematic issue is, many hospitals are yet to implement security in the notion of data transmission. Information of a consumer involves doctors' reports, scan reports and other vital reports. Recently, a security breach has came to a notice giving a scope to doubt the technique of concealing the data securely in a Cloud Computing.

ISSN: 2005-4238 IJAST

Similar devices procured from different manufacturers who have produced the devices with varied standards is also a potential issue to be addressed as inconsistent standards are always troublesome to maintain a smooth transmission in the terms of bandwidth and speed that change simultaneously along with standards. An immediate effect of inconsistent bandwidth is that, coherent data is no more a guaranteed luxury. With changing frequencies there is an undeniable effect on the consumption of power and the efficiency is in direct risk. Cost and agility in system design are other weights that may have to be considered as they also change with changing standards. Some common situations, threats and challenges are shown at Fig. 3. Internet security and privacy are those which are being misused and may needs to be debated, when the resources are available. Further, there may be situations when few consume the whole resources and other authenticated consumers are denied which is commonly referred as resource hacking in the literature. Authorisation is passing the information to the mentioned user without having any encryption performed on the information from the data to the consumer. Approving authorisation to consumers to look at the information and utilize resources can lead to threat due to cyber criminals who can decrypt the data to get access to the information without any requirement of authorization.

Cryptography is the science of encryption and decryption techniques in order to keep the data confidential. It is like making the data not understandable till there is an authentication. Algorithms for encryption and decryption are undergoing periodical changes and this is one of the vulnerable zone in the systems where hackers could target upon. Static firmware which is incapable of auto updating may face security issues since they cannot cope with newer encryption algorithms. Notable point is that there are only small percent devices which can update firmware automatically. Henceforth, with time older generation devices are likely to be more prone for hacking or corrupted. Whatsapp, removing devices with ios 8 or older from its services are also a targets for hackers.

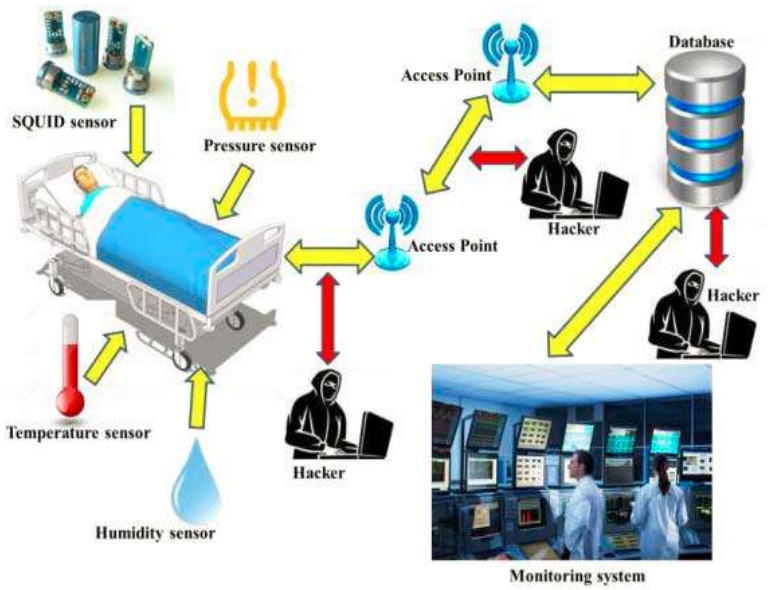

Figure 3 Common Scenario in Internet of Things based Healthcare and the Potential Threats and Challenges [10].

To gain an understanding about real life mining and how it could enhance the healthcare industry, some examples are provided:

a) Fat scales which can be used to diagnosing the persons weight, water, fats, cholestrol, body balance related issues are in existence with the help of IOT.

ISSN: 2005-4238 IJAST

b) We can take a record of how frequent a person will be unable to meet his schedules set in his mobile and predict his level of distractions henceforth advising meditation sessions for improving mental peace and decreasing stress.

Limitations and Challenges of Data Analytics in Healthcare Industry

As everything has its merits and demerits similarly in data analytics, there exist few challenges to implement data sciences in health care sector:

a) The medical data providers (hospitals, companies, druggists etc.) are in diverse standards. Those institutions have data in the formats viable to their usage rather than universal standards. Usage of such heterogeneous data costs those organizations to share data warehouses for making the data readable in some common formatting mode and someone has to manage those situations. Further, there is an immediate upsurge in finance required for getting such luxury of data warehousing.
b) Data acquired is not necessarily a proper design of real life happenings and therefore, it has to be verified as well as processed in order to assure quality for data analysts. This suggests that quality is also a limitation in this field of heathcare.
c) Big investments are necessary in organisations to make such plans work. Organisations have to hire data scientists, engineers, analysts, designers as well as pay medical organisations to convince them to invest in data analytics.
d) Data science is nothing but complex mathematics using computers. Thus talent is required and organisations have to pay heavily to hire such talent.
e) Further, there is always a possibility of error due to improper collection of data.

To sum up, Internet of things based healthcare with respect to Big Data analytics have several critical issues to address and the obstacles in implementing data science are discussed in brief in this section. Limitations and challenges listed above are just a sample.

VII. CONCLUSIONS AND FUTURE ENHANCEMENTS

We have introduced the concept of data and its year by year evolution. We have also covered the importance of data and its role in retailing, health care, banking, and others applications. We found that e-healthcare is most important area to focus in research community. In particular, Artificial Intelligence is developing at an alarming pace, there will be technologies in existence which can indicate one selves future ailments. We have covered the existing tools available to refine data and also the current issues exist in the healthcare field. Then in this work, we addressed several issues and challenges with respect to existing tools and algorithms to refine big data, i.e., to predict disease at earlier stages. In near future, computing power will raises to a state of pure excellence in treatment and diagnosis of patients with surgical precision of humans. We need answer to questions like “tracking the evolutionary progress on how data has been collected, stored, managed and analysed”? Further, we discuss several suggestions to overcome these issues. We have concluded that overall view of data science and data mining for acquiring, analysing and maintaining healthcare info from multiple sources is a necessity of future. Additionally, enhancing new technologies and modern tools takes our attention to build a robust and secure health system which satisfies all involved parts. Hence in conclusion, we need a new method/tool/ algorithm to refine sensitive data in e-healthcare. So we need to propose a predictive model, which can be used to train the dataset in a well-defined manner to retrieve useful information/ in a cost effective manner. Later, this information can help to even prevent some spreading diseases. Further, Internet of Things can play a major role in completing the dream of autonomous vehicles for future generation, but still there is a long way to go to achieve such expectations. Hence, whoever (researchers from around the

ISSN: 2005-4238 IJAST

globe) is working in this area/ interested to do research in this area, are invited to do their work in the respective field.

Authors' contributions

Both A. Mohan Krishna and Amit Kumar Tyagi have contributed equally. Amit Kumar Tyagi has approved the final manuscript.

REFERENCES

[1] Raghunath Nambiar, Ruchie Bhardwaj, Adhiraaj Sethi and Rajesh Vargheese, "A Look at Challenges and Opportunities of Big Data Analytics in Healthcare", IEEE International Conference on Big Data, **(2013),** October 6-9.
[2] J.Archenaa and E.A.Mary Anita, "A Survey of Big Data Analytics in Healthcare and Government", International Symposium on Big data and Cloud Computing, Procedia Computer Science, Elsevier, **(2015),** pp 408-413.
[3] Raghupathi Wullianallur and Raghupathi Viju, "Big data analytics in healthcare: promise and potential", Health Information Science and Systems, Vol. 2, **(2014).**
[4] Peter Groves, Basel Kayyali ,David Knott and Steve Van Kuiken, "The 'big data' revolution in healthcare: Accelerating value and innovation", Center for US Health System Reform Business Technology Office, McKinsey & Company, **(2013).** Available at https://www.ghdonline.org/uploads/Big_Data_Revolution_in_health_care_2013_McKinsey_Report.pdf
[5] Liang Hong, Mengui Luo, Ruixue Wang, Peixin Lu, Wei Lu and Long Lu, "Big data in Health care: Application and Challenges", Data and Information Management, Vol.2, Issue.3**(2018),** pp:175-197
[6] Tempest A van Schaik, Nadezda V.Kuveleskaya, Elena Protopapas, Hamzawahid and Fiona G.G.Neilson, "The Need to redefine genomic data sharing: A focus on data accessibility", Applied & Translational Genomics, Vol. 3 Issue.4 Elsevier, **(2014),** pp: 100-104
[7] Hiba Asri, Hajar Mousannif, Hasan Al Moatassime and Thomas Noel, "Big data in Healthcare: Challenges and oppourtunities", IEEE International conference on Clod Technologies and Applications, **(2015),** June 2-4.
[8] Jingquan Li, "Open Medical Big Data and Open Consent and their Impact on Privacy", IEEE International Congress on Big Data, **(2017),** June 25-30.
[9] R.Sathyavathi, "A survey: Big data analytics in Healthcare system", Contemporary Engineering Sciences, Vol. 8, no. 3, **(2015),** pp: 121 -125.
[10] Sharath Anand and Sudhir R Routray, "Issues and challenges in healthcare narrowband IOT", IEEE International conference on Inventive communication and computational Technologis **(2017)**, March 10-11.
[11] https://www.cso.com.au/article/575407/internet-things-iot-threats-countermeasures/
[12] José Moura and Carlos Serrão, "Security and Privacy Issues of Big Data" Handbook of Research on Trends and future directions in Big Data and web Intelligence, **(2016),** pp: 20-52.
[13] Veerendra G.G, "Hacking Internet of Things (IoT), A Case Study on DTH Vulnerabilities", SecPod Technologies, available at https://www.secpod.com/resource/whitepapers/Hacking-IoT-A-Case-Study-on-Tata-Sky-DTH-Vulnerabilities.pdf.
[14] https://www.senetas.com/network-traffic-flow-analysis-protection-enhances-encryption/
[15] Mark Van Rijmenam, " A short history of Big Data", available at https://datafloq.com/read/big-data-history/239
[16] Amir Gandomi and Murtaza Haider, "Beyond the hype: Big data concepts, methods and analytics", International Journal of Information Management, Volume 35, Issue 2, **(2015),** pp: 137-144.

ISSN: 2005-4238 IJAST

[17] Amit Kumar Tyagi and G, Rekha, "Machine Learning with Big Data", Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), **(2019),** available at SSRN: https://ssrn.com/abstract=3356269

[18] Amine Rghioui and Oumnad Abdelmajid, "Challenges and Opportunities of Internet of Things in Healthcare", International Journal of Electrical and Computer Engineering (IJECE), Vol 8, No.5 **(2018),** .pp2753-2761.

ISSN: 2005-4238 IJAST