Security Challenges of Big Data Implementation in Health Care Systems

Dr. Senthil Kumar M  
Associate Professor  
Department of Computer Science & Engineering  
Valliammai Engineering College

Dr. Chidambararajan. B  
Professor & Principal  
Department of Computer Science & Engineering  
Valliammai Engineering College

Abstract

As technological improvement in data management techniques, Big Data tend to handle much more volumes of data. There might be higher beneficial in managing data , on the other side there is a possibility of having much more challenges in providing security for that data from the intruders. The main aim of this paper is to convey how big data works in the health care system and security related challenges in the implementation. This paper may be helpful to the researchers to find the challenges in the big data and it also gives way to find the solution for that security issues.

Keywords: Data environment, Metadata, challenges, Access Control, Health care system

I. BACKGROUND

Recent advances in rapid growing technology a huge amount of data has been generated and stored in databases, there may be a presence of various issues and challenges in Big data environment. Big data a term denotes large volume of data that can be stored, analyzed to obtain the required output. The major challenges in big data are Capturing of data, Curation, Storage, Searching, Sharing, transfer, Analysis and Presentation [1].

The main characteristics of big data are volume, velocity, variety. Big data architecture for capturing the data from various data streams as well as analytic process is represented in fig 1.

The data streams from various sources are obtained. They are tend to store on data storage infrastructure where it can be structured or unstructured and analytic process are done according to the privileged rights assigned to particular users (such as Data owners, technical analyst, business analyst).

Fig. 1: Architecture of Big data

II. CHALLENGES IN BIG DATA

The various challenges in big data are as follows [4]. They are

- Distributed frameworks are those in which huge data processing are paralyzed and establish a communication across various systems
- Non-relational data stores in that retrieval of data is little tedious to implement a privacy related features
- Storage in which volume of data can be exploded frequently.
- Endpoints are those which need to validate the authentication process.
− Real time security helps to ensure privacy during the end to end interaction.
− Data mining solutions.
− Access controls is one of the factor that need to be consider for assigning the privileges. It helps to limit the users according to their assigned rights.
− Granular auditing provides a statistical data regarding the past events and provides efficiency in security related issues [6].
− Data provenance contains the metadata (data about data) in determining the data source, privileged rights and analyzing process.

III. BEST PRACTICES

The following are the guidelines for good practices to be managed in Big data. They are
− Creating dimensions in data store, that has to be divided into dimensions and facts.
− Dimensions must have a durable surrogate keys which can’t be changed once assigned.
− Structured as well as unstructured data needs to be integrated and analyzed together.
− Technology [3] can implement different forms of data or it needs to be generalized.
− Data privacy needs to be ensured.

IV. REVOLUTION IN BIG DATA

The various sectors where the evolution of big data technologies are Banking and securities, communication, media & entertainment, health care providers, Education, Manufacturing & natural resources, government, Insurance, retail & wholesale trade, Transportation, energy utilities etc. In banking sector it has Security Exchange Commission where financial markets are monitored by network analytics and natural language processors [5]. In Government side, Food and Drug Administration (FDA) helps to detect and study patterns of food related to illness and diseases, faster response to treatments. The main focuses on health care system how big data uses in real world applications and some challenges faced in real time systems [2]. In Fig -2, the architecture of health care systems are represented in which the data source are from various areas such as hospital, public health, research institutions, patient, network operator and management process are established at data storage infrastructure. It includes three main layers of implementation. They are
− Data collection layer
− Data management layer
− Application service layer

A. Data Collection layer

In this layer, the data are collected from various sources such as health, patient etc. It can be stored in either structured, semi structured or an unstructured data.

B. Data management Layer

It includes various activities in processing of data. They are
− Raw data obtained from data collection layer.
− Data Transformation is the process of converting the obtained raw data into an standard format.
− Distributed file storage it has a data which are in those specified standard.
− Data parallel computing allows multiple users to process in a parallel manner.
V. CHALLENGES IN HEALTHCARE SYSTEMS

The fundamental challenge associated with big data is how to describe, investigating, stimulate and manage the dynamics of various types of data. The vital point that drives this data analytical process is that there is a need to manage the large information. The issues and challenges in health care system are mentioned below. They are

- Multi source Heterogeneous Data management includes data obtained from various sources needs to be managed which are in different format.
- Diversified data analysis includes process of analyzing processing those data into standard format
- Application service platform provides a standard interface to establish communication. Those interfaces are API and other protocols.

VI. CONCLUSION

Big data applications are growing at exponential rate and related issues are growing accordingly in various disciplines such as medicine, Aerospace, banking sector, health care, computer science and IT, there is a growing need to encompass this ever demanding field. Managing heterogeneous data in the health care system is needs more techniques and also requires security parameters to safeguard the data. With the continuous improvement in technologies and accumulation of data at an under estimated rate, there will be continuous challenge to sort out the problems of technology development. This paper provides the platform to identify the challenges in implementing security features to the data.

REFERENCES

[2] Yin Zhang , Mohammad Mehedi Hassan ,”Healthcare Cyber physical system assisted by cloud and big data” on IEEE SYSTEMS JOURNAL