An Outlook on India's Healthcare System with a Medical Case Study and Review on Big Data and its Importance in Healthcare

Saikrishna Depuru[1]*, Dr. R. Raj Kumar[2]

[1]Research Scholar, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

[2]Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

Abstract: *In the present day world, data management has become an essential and integral part of every organization and the volume of data is getting multiplied day by day at an alarming rate. Likewise, the growth of Indian healthcare industry is at an enormous pace and there is a revolutionary increase in the size of healthcare data. The objective of the paper is to discuss about the existing health care model in India along with a case study showing the significance of healthcare data management and analytics using Bigdata. An on-premise Hadoop based healthcare data management system is proposed showcasing the importance of Big Data Analytics and the way it could help the health care industry to grow and provide the quality of service to patients.*

Keywords: Healthcare, Data Management, Patients, Bigdata Analytics, Hadoop.

1. Introduction

Health plays a major role in human life and it has become a global social objective. Health is required for understanding of basic human requirements and to gain the status of improved quality of life. Contrary to the above background, the India's healthcare system is encountering serious issues in affording good health care to its citizens. Despite of all the health problems in the country, India's public contribution on health is remarkably low because of many challenges [1]. Due to the quantity and quality gap of service in the public healthcare system, patients are moving to private healthcare providers to have a better quality of treatment. These private healthcare is constituting a larger part of the healthcare system in India and are presently treating 80 percent of outpatients and 60 percent of inpatients [1,2].

Along with the hospitals ,the healthcare industry has become an the integration of various fields like pharmaceutical, Medical Insurance, biotechnology, related life sciences etc [3].Hence the Indian healthcare industry has witnessed impressive growth in the past few years and even continuing its growth and transforming in a rapid phase. Apart from that, the other driving factors for this growth could be increasing need for good health care facilities, health policies and advancing awareness about health in the public.

The Indian healthcare sector was predicted to be US\$ 40 billion during year 2010 and it is forecasted to reach US\$ 280 billion by year 2020 [4].According to the analysis provided by 'India Brand Equity Foundation' the growth trend of the healthcare industry is shown in Figure 1.
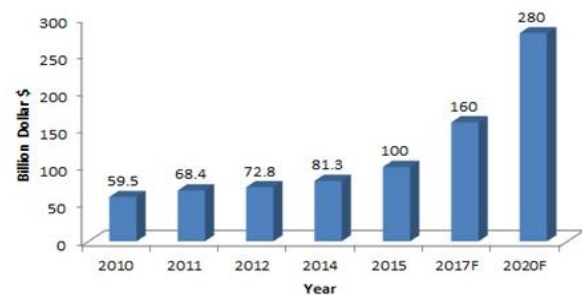

Figure 1: Growth Trend of India's Healthcare Sector.

Recent report published by 'Indian Hospital Services Market Outlook' shows that the investment done by private sector players in the Indian hospital industry played a major role in the growth of the healthcare sector and at present holds a major share of 75 -80 percent of the entire health care market [5]. Of the total health care revenues, hospitals account for 71 per cent and the rest is from different areas [6].
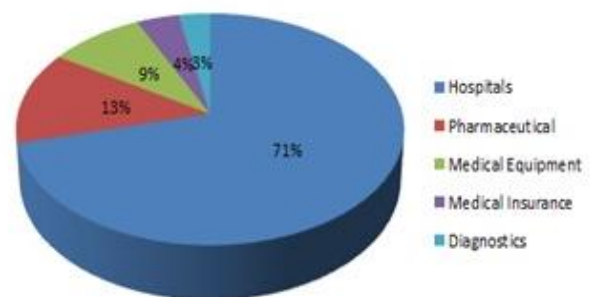

Figure 2: Market breakup by revenues of total health-care revenues in India

Paper ID: NOV162856

1486

1.1 Healthcare IT

Information technology (IT) is playing a crucial role in transforming care delivery and helping in raising the quality, safety and performance of healthcare delivery. Furthermore, the information could also be used to institutionalize the delivery process and can be further leveraged to result in new methods of healthcare.

The involvement of IT will help health care industry to evolve with innovative and adept healthcare infrastructure required for urban and rural community. IT resolutions are on demand with a high priority to improve resource optimization, schedule doctors/therapist, pharmaceutical orders, laboratory test orders etc. The progress in healthcare IT sector will also help hospitals in reinforcing the efficiency of their organizational activities hence accelerating to reduce patient waiting time, prevent medical errors and enhancing success rate in treatment [7]. But India's healthcare technology investments are quiet small related to the overall population in country.

2. Data Management

Along with many issues the major problem faced in healthcare industry is data maintenance and management. Hence improper and inaccurate data may adversely influence the quality of an individual's healthcare. Various studies have exhibited that the confined access to patient's information during decision-making and the ineffective storage, exchange and analysis of information among doctors are proximal source of medical flaws in healthcare [8].As data is being generated day by day in different formats like structured ,semi structured and unstructured , there is a necessity to store, analyze and act on data in the health care industry to bring out the good insights from it to enhance patient safety and improve quality of service.

Computerization of health information and the scope of organizational storage, exchange and analyzing huge amount of data by maintaining integrity and completeness should be taken care and this will be a paramount .Accomplishing this goal is dependent on the capability to store and perform analysis on large volumes of data.

3. Existing Healthcare Model and its disadvantages

As of today, the Indian healthcare is mostly a "sick-care" industry represented by many diseases, serviced by an inadequate and highly fragmented provider infrastructure, catering to a population that has to spend out of its pocket for most of its healthcare needs, but still does not have an assurance of receiving quality care. The existing healthcare model in India need to be improved and should move a step ahead in terms of technology to provide good health to the patients.
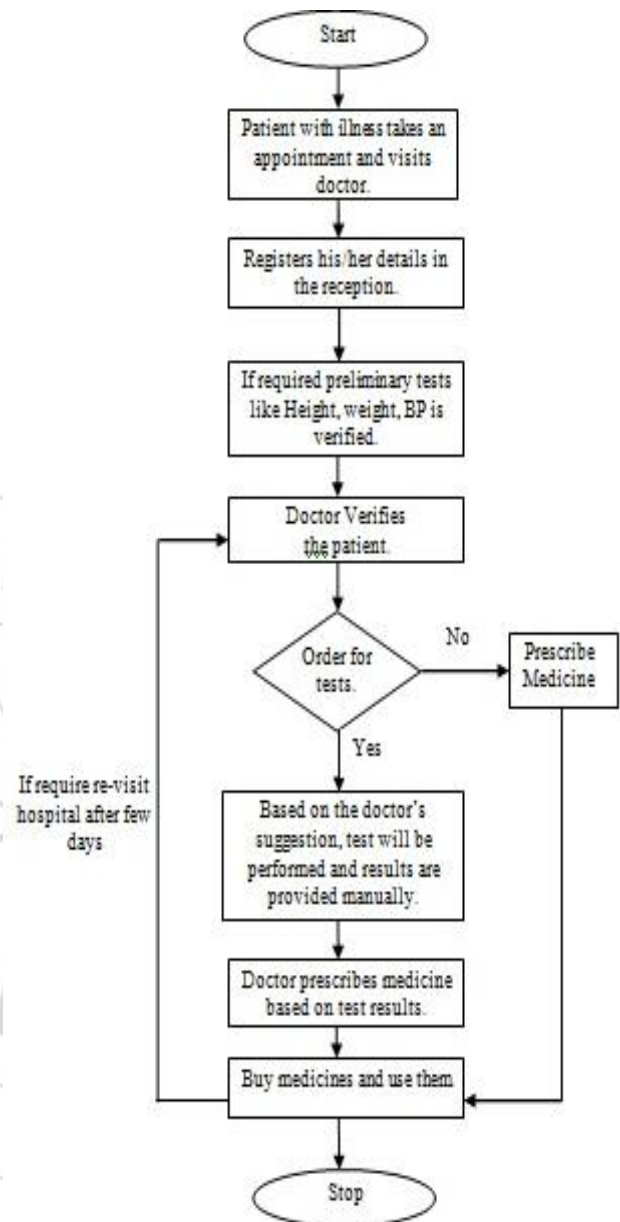

Figure 3: Existing healthcare model in India

As shown above, patients maintaining paper based prescriptions or results at different stages of treatment is a traditional method and has many disadvantages especially when the patient is treated for longer period of time.
Manual medical records causes many draw backs like

- **Expensiveness of manual records**: One of the most expensive disadvantages of the paper record is repetitive testing might be required when a old reports or test results are lost.
- **Accessibility:** In a critical situation, if a medical record needs to be accessed by an expert at a different location, the medical records cannot be accessed immediately.

Paper ID: NOV162856   1487

- **Downturn in work rate**: The time taken to search the records and delivering them is very time consuming process and will affect the productivity.
- **Maintenance of Good Quality records:** There are limitations to the physical quality of the paper record.

4. A Medical Case Study

In order to get into more details of prevailing healthcare model in various hospitals, especially hospitals treating chronic disease like cancer, diabetes etc and to showcase the importance of data management and analytics, we have conducted a medical case study to get the details from a cancer patient and the way he was diagnosed and treated.

Table 1: Details of patient

A Cancer Patient Case Study	
Name	Subramanyam
Age	50
Location	Nellore,India
Cancer type	Chronic Lymphocytic Leukaemia.
Year of Diagnosis	2014
Hospital	Cancer Institute (WIA) ,Chennai,India.
Method of treatment	Chemotherapy
Number of Chemo Cycles	6
Treatment - time period	5-6 months
Tests performed during treatment	Blood Tests, Bone Marrow, Scanning(MRI & PET)
Frequency of visit after treatment	2 months

4.1 Diagnosis & Clinical Trials

Mr. Subramanyam was diagnosed with cancer in August 2014 after presenting symptoms of unexpected weight loss and swollen lymph nodes across various body parts. After the diagnosis, his oncologist has decided to treat him through chemotherapy in 6 intervals and the clinical trial experience was positive. He was hospitalized only for the first chemo cycle and for the remaining, he underwent chemotherapy in the day care. During the process of treatment he has under gone various tests like bone marrow, blood test and PET scan. After the successful treatment of cancer, he still visits doctor every 2 months for regular health checkup. He still persists with some undiagnosed health issues after the cancer treatment.

Table 2: Questionnaire to patient

Questions	Answer
Whether all the test reports are stored on a computer?	Partially
How frequently the blood tests were performed during treatment?	Every 15-20 days
How many times a blood test was performed before, after and during treatment (in the period of 1 year)?	15 times
How many times a bone marrow test was performed before and after treatment?	3 times
How many times a Scanning was performed before and after treatment?	3 times
Do you feel, whether there was any delay in appointment of doctor due to maintenance and access to manual health records?	Yes
Do you feel there was any ambiguity in recognising the symptoms before and after cancer treatment?	Yes

From the above case study we understand that , lot of data is being generated from clinical reports, discharge summaries, medicines and many more which can be stored and analyzed. As this data has a lot of useful information and with the evolution of numerous planning and analysis techniques, big data is becoming a foundation for an effective information production and decision making advancement [9]. Hence, the application of big data analytics in the healthcare will help in

- Monitoring the health conditions, based on the consolidated test results.
- Extracting the hidden and unknown facts from the existing health records of patient who have under gone treatment for similar kind of disease.
- Provides an opportunity to doctors to look and analyze at different methods/approaches of treatment, where a patient has already been treated for a similar case.
- Predictive analysis can be performed based on the present and past data.

And also looking at the existing healthcare model and its disadvantages, there is a great necessary in the healthcare to perform data analytics, especially while treating the chronic disease like cancer, diabetes etc.

Along with this, we have performed a survey with set of doctors to know importance and necessity of EHR and data analytics in treating chronic diseases. And we could find the response was quite favorable to data analytics in healthcare. So in this paper the advantage of bigdata - hadoop and big data based healthcare data management models/applications are discussed.
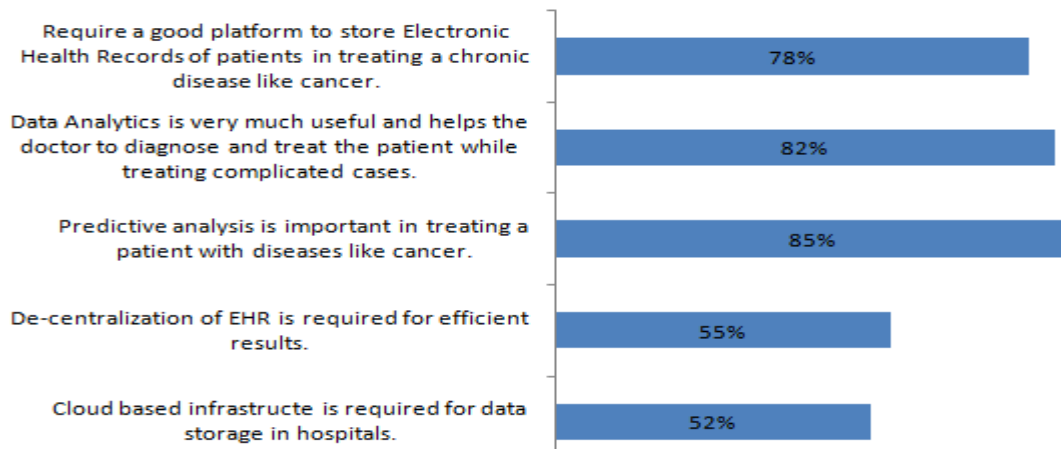
Paper ID: NOV162856

1488

Figure 4: Survey on importance of EHR and data analytical in treating chronic diseases

5. Hadoop based healthcare data Analytics

Considering different issues in maintaining manual medical records, the patient's information and the medical information can be saved in the form of EHR. As data is loosely structured [10] and big in size, new method of data storage is required.

5.1 Big Data Vs Traditional Databases

Big data is a word that represents the huge volume of data comprising of both structured and unstructured data. The common consent of the day is that there are particular attributes that describe big data, these are labeled as five V's: volume, variety, velocity, veracity and value [11]. In most of business scenarios the volume of data is huge or it multiplies fast or it outpaces the prevailing processing capacity. Big data can be Terabytes, Petabytes of data consisting of billions to trillions of records from different sources (e.g. Sales, social media, Web, customer contact center, mobile data etc) [12]. As data is usually unstructured which is frequently incomplete and unavailable, the traditional RDBMS cannot handle all kinds of data from healthcare, which are organized to handle mostly structured data. Whereas BigData-Hadoop can handle various types of data.

5.2 Hadoop

The term big data, specifically when used by vendors, may refer to the technology, including tools and processes that an organization requires handling huge amount of data and storage facilities [13]. The Apache Hadoop was born out of a need to process an torrent of big data and this software library is a framework that provide scope for the distributed processing of massive data sets across group(clusters) of computers by making use of simple programming models. Scaling up from single server to multiple machines and each offering local computation and storage is possible in system [14].Hadoop has many additional advantages and different properties compared to existing traditional relational databases [15].

Table 3: Difference between RDBMS and Hadoop

Factors/Elements	RDBMS	Hadoop
Data Size	Gigabytes to Terabytes	Petabytes or Exabytes
Data Integrity	High(referential,typed)	Low
Data Schema	Static	Dynamic
Data Structure	Structured	Structured and Unstructured
Data granularity	Stores transformed, aggregated data	Stores huge volume of highly granular data.
Data Normalization	Required	Not Required
Analysis type	Exploratory analysis to uncover value in the data	Operational analysis of what was uncovered
Query Response Time	can be near immediate	Has latency

6. Review on Big data in Healthcare

Approach 1: The authors [16] has provided information on promise and potential of big data analytics in healthcare by explaining the benefits, challenges of bigdata.

Results: The paper provides a synopsis of big data analytics for healthcare researchers and practitioners and an outline of big data analytics in healthcare methodology .The conceptual architecture consisting details about Data sources, Data transformation, Big Data Tools (Hadoop,MapReduce,Pig Hive,HBase etc) and analytical applications(Queries,Reports, OLAP, DataMining) are proposed here [16].

Approach 2: In paper [17], an analytical model was proposed for an extended version of Artemis system which is being arranged at SickKids hospital in Toronto. The Artemis Cloud is designed in such a way that it supports acquisition storage of physiological data and clinical information for the purpose of visualization, realtime analytics and reflective analysis.

1489

Results: Using the performance model, valuable performance metrics such as mean number of patients in NICU, mean patient residence time, mean number of required medical algorithms and blocking probability were characterized and discussed. Based on pilot project at SickKids, the amount of required storage, memory and computing power for analytics and real time components respectively were determined [17].

Approach 3: A framework for Healthcare Information Systems based on big data analytics in mobile cloud computing environment was introduced. This furnishes a high level of availability, integration, interoperability and sharing of health records among healthcare providers, patients, and practitioners [18].

Results: The paper provides a broad overview of proposed framework where Electronic Medical Records patients distributed across different Care Delivery Organizations are unified and stored in the cloud storage area which creates an Electronic Health Records for each patient.

Approach 4: This paper [19] presented the recent research using Big Data tools and approaches for the analysis of Health Informatics data gathered at multiple levels like molecular, tissue, patient, and population levels. In addition to gathering

data at different levels, multiple levels of questions were addressed: human-scale biology, clinical-scale, and epidemic-scale [19].

Results: The methods of analyzing data at different level all of these levels in order to answer the ever-growing list of medical questions on all of these levels. The studies covered here seek to answer questions on all four levels from data through analysis of data from the molecular, tissue, patient and population levels.

7. Hadoop based On-Premise Data Management System

Considering the advantages of Hadoop over the traditional RDBMS and from the proposed models and study on Big data analytics , a hadoop based on-premise framework for healthcare data management and analytics is proposed .And, more over from the real-time case study we could understand there is a great necessity for healthcare data management and analytics in the big hospitals where the data inflow is high. So this Big data Hadoop environment can be created in a data-driven hospitals providing the scope for massive data storage and analytics.
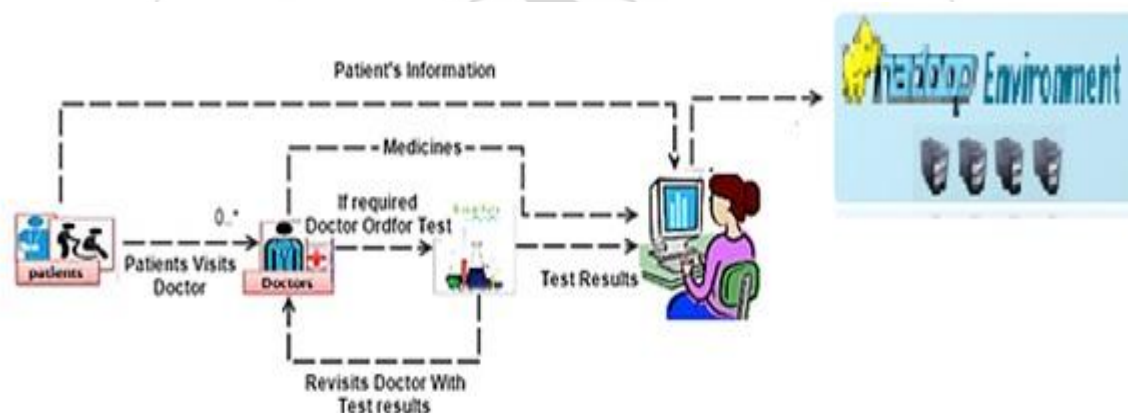

Figure 5: On-premise healthcare data management using Hadoop

The above proposed method has many advantages over the traditional approach.
- Storage and access to huge amount structured and unstructured data.
- Sharing and accessing information.
- Data Quality is maintained.
- Increases the productivity due to faster data transfer and providing bigger and better opportunities for data analytics.

7.1 Goal of Big data analytics in healthcare

The goal of big data analytics in healthcare [20] is to

- Make use of the huge amounts of data and provide right intervention to the right patient at the right time.
- Personalized care to the patient.
- Potentially benefit all the components of a healthcare system i.e. provider, payer, patient and management.

According to a survey from MeriTalk of 150 federal IT and business executives from healthcare agencies the benefits of big data in healthcare are show below [21].
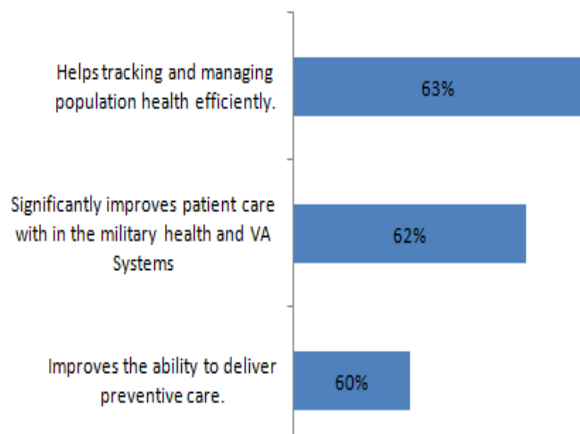
Paper ID: NOV162856

1490

Figure 6: Benefits on Big data in healthcare

Similarly in another perspective the top goals for using analytics, according to a survey of 40 hospitals and 30 insurers according to Information Week, "Healthcare Organizations Go Big For Analytics are shown below [22].
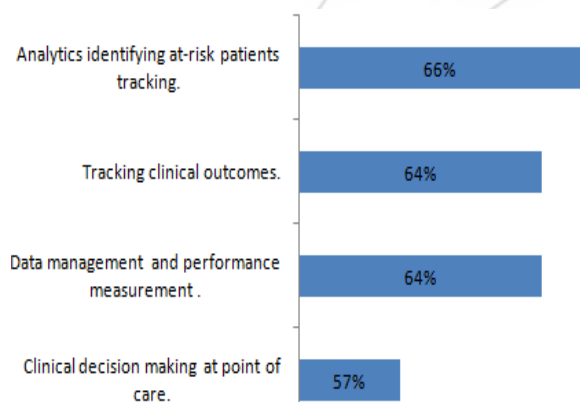

Figure 7: Survey Results on Hospital's goals to use big data analytics.

8. Conclusion

The paper presents a brief outlook on India's healthcare system and the existing healthcare model and its disadvantages. A case study was performed on experiences of a person who was suffering from cancer and the details, like the way he was diagnosed and treated was studied in this paper. And from the study, the production of data and its necessity to store and analyze was recognized and along with this, various proposed approaches studied towards big data analytics in healthcare was reviewed. From the real-time case study and understanding of different approaches the importance of bigdata analytics are discussed and an on-premise healthcare model was proposed to store and analyze data specific to that hospital.

References

[1] Vikas Bajpai. The Challenges Confronting Public Hospitals in India, Their Origins, and Possible Solutions Advances in Public Health, Volume 2014 (2014), Article ID 898502, 27 pages.

[2] Susmita Chatterjee,Carol Levin,Ramanan Laxminarayan. Unit Cost of Medical Services at Different Hospitals in India. PLOS ONE. July 23, 2013.

[3] Kunjana Malik.Outlook on the Private Equity Potential in India's Health Care Sector Based on Lessons from US. International Journal of Multidisciplinary Consortium, Volume -1, Issue – 1,June 2014.

[4] Naresh Vashist, Puneet Jain. Private Sector In Indian Hospital Industry-A Review. International Journal of Transformations in Business Management, 2013, Vol. No. 2, Issue No. 3, Jan-Mar.

[5] G. Nagarajan,Dr. J. Khaja Sheriff. Novel Healthcare (Hospitals) – Creation An Indian Perspective. International Journal of Marketing, Financial Services & Management Research.Vol.2, No. 4, April (2013).

[6] Dr. Md. Izhar Alam, Dr. Jameel Ahamad Khader. An Analysis of Foreign Direct Investment Inflows in Healthcare Sector In India. Pezzottaite Journals, Volume 4, Number 2, April – June' 2015.

[7] Sharma Kalpa. Health IT in Indian Healthcare System: A New Initiative, Research Journal of Recent Sciences, ISSN 2277-2502, Vol. 1(6), 83-86, June (2012).

[8] R.Anand.,Dr.S,K.Srivatsa.An Implementation of Health Care Industry through Cloud Computing Technology. International Journal of Emerging Technologies in Computational and Applied Sciences,2014.

[9] Keon Won Kim , Won Joon Park and Seong Taek Park , A Study on Plan to Improve Illegal Parking using big Data. Indian Journal of Science and Technology, Vol 8(21), September 2015.

[10] Abdul Raheem Syed , Kumar Gillela , Dr. C. Venugopal . The Future Revolution on Big Data. International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 6, June 2013.

[11] Demchenko, Y, de Laat, C, Membrey, P . Defining architecture components of the Big Data Ecosystem. Collaboration Technologies and Systems (CTS), IEEE, 104 - 112 ,19-23 May 2014.

[12] Ravikumar H. Roogi. Big Data Solution by Divide and Conquer technique in Parallel Distribution System using Cloud Computing. Oriental Journal Of Computer Science & Technology, Vol. 8, No. (1), 09-12,April 2015.

[13] G. Yogaraj , A. Arumuga Arun.Mining High Dimensional Data Sets Using Big Data.International Journal of Advanced Research in Computer Science and Software Engineering. Volume 5, Issue 2, February 2015.

Paper ID: NOV162856

1491

[14] Vidyullatha Pellakuri , Dr.D. Rajeswara Rao. Hadoop Mapreduce Framework in Big Data Analytics. International Journal of Computer Trends and Technology ,volume 8 number 3,Feb 2014.

[15] D. Rajasekar , C. Dhanamani , S. K. Sandhya. A Survey on Big Data Concepts and Tools.International Journal of Emerging Technology and Advanced Engineering. Volume 5, Issue 2, February 2015.

[16] Wullianallur Raghupathi, Viju Raghupathi.Big data analytics in healthcare: promise and Potential.Health Information Science and Systems 2014.

[17] Hamzeh Khazaei,Carolyn McGregor, Mikael Eklun d, Khalil El-Khatib & Anirudh Thommandram .Toward a Big Data Healthcare Analytics System. A Mathematical Modeling Perspective : Services (SERVICES), 2014 IEEE World Congress,208-215.

[18] Ahmed E. Youssef .A Framework For Secure Health care Systems Based On Big Data Analytics In Mobile Cloud Computing Environments, International Journal of Ambient Systems and Applications (IJASA) Vol.2, No.2, June 2014.

[19] Matthew Herland,Taghi M Khoshgoftaar & Randall Wald.A review of data mining using big data in health informatics, Journal of Big Data 2014.

[20] Jimeng Sun ; Chandan K. Reddy ,' Big Data Analytics for Healthcare' ,Tutorial presentation at the SIAM International Conference on Data Mining, Austin, TX, 2013.

[21] Dr. Amit Dang , Dr. Shilpa Mendon.The Value of Big Data in Clinical Decision Making. International Journal of Computer Science and Information Technologies, Vol. 6 (4) , 2015, 3830-3835.

[22] How Big Data Impacts Healthcare , Harvard Business Review, Harvard business publishing ,2014.