A NOVEL HEALTHCARE MODEL TO EXPLORE CLOUD DATA SPACE DIGITALLY FOR HEALTH IMPROVEMENT IN RURAL AREAS

Dr. B. Mukunthan (PhD)
Research Advisor, PG Research Dept. of Computer Science
Jairams Arts and Science College, Karur-3
Bharathidasan University, Trichirappalli -24
Tamilnadu, India

Mrs. B. Radha (MCA)
PG Research Dept. of Computer Science, Jairams Arts and Science College, Karur-3, Bharathidasan University, Trichirappalli -24, Tamilnadu, India

Abstract: Qualitative health care services are considered as a basic need for every citizen irrespective of age, gender and culture. The Health Care system in developing countries still relies on paper medical records. Information that is digitized is usually not portable; preventing information sharing amongst the different health care resources. The Health Care system faces substantial challenges with growth inhibitors such as rapidly growing population, changing disease profile and re-emerging diseases, multilayered healthcare system or landscape, lack of access infrastructure, scarcity of manpower that is doctors, nurses and paramedics, extremely low public expenditure on health and its inefficiencies. The aim of this paper is to create an improvised public health care system for developing countries with efficient big data analytics to delve bioinformatics data taking into account from various resources including medical records, hospitals, research groups, insurance companies and public and private sector health organizations. The proposed Health Care model facilitates collaboration and to coordinate care between patients and physicians and amongst medical community possible. It also provides an infrastructure that allows hospitals, research groups, insurance companies and public and private sector health organizations to tap improved computing resources at lower initial investments. Also the cloud environment will reduce the obstacle for innovation and modernization of health information technology systems and applications. This paper also includes guidance and strategies that helps to build an autonomous, digital and efficient health care system to benefit the rural population utmost and nation as a whole.

Keywords: Healthcare, Internet of Things, Big data, Data Mining, Bioinformatics, Neural Networks, Cloud computing, Biological Database.

I. INTRODUCTION

With ceaseless augmenting of population, technology, foods, air pollution, etc. the huge issue of treating and preventing disease among nation requires the study and a careful strategic approach. This is really essential to keep the nation of worldwide safe from maladies without cure. For Instance rural India subjects to 50-70% of non communicable disease per 1000 cases with 50 million cardiac patients; 63 million diabetic patients; 1.2 million cancer cases every year; 0.9 million stroke cases every year; 1,75,000 renal transplants.

The rural population contributes 72% i.e. 742 Million with 285 million population urban shares 28% of the whole. In rural and urban the hospital% is 31% and 69% ;hospital bed% is 20% and 80%; Doctor% is 8% and 92%; spurious pharma is 75 to 80% and 20 to 25% respectively moreover the doctor/100,000 people is 05 in rural and 50 in urban which clearly indicates health infrastructure in India is well below the WHO recommendations. The National Family Health Survey [NFHS] report for the year 2015-16 on the chronic diseases asthma and diabetes is shown in Figure 1 and Figure 2 respectively. Discovering of relevant doctors is a problem for even patients living in tier-1 cities which are still a nightmare for those in rural areas.

II. DATA MINING

Data mining is often defined as finding hidden information in a database and is also known as analysis steps of knowledge discovery in database. The main goal of data mining process is to extract information from a data set and transform it into an understandable structure for further use. The huge amount of data collected in data mining [1] is termed as big data. Big data is a term that describes at least three separate, but interrelated trends: capturing and managing lots of information, working with many new types of data and exploiting these masses of information and new data types.

Figure 1 National Family Health Survey-NFHS Report on Asthma
This data mining covers the whole part of the education disciplines such biological data, mathematical, physical, medicinal etc. Study that marvelous organism, the human system, and the laws by which it is governed is one of the earthshaking area of economy; as such for a society to realize its dream and vision of sustainable growth, it is very important to know that every careless action, any abuse put upon the wonderful mechanism, by disregarding his specified condition the future of nation will suffer. For this precious life there is a saying among society which says that “health is wealth”.

Management issues

The biggest data management is collection of large volumes of structured, semi structured and unstructured data from data heterogeneity and variety arise as a result of linking the diverse range of biomedical data sources available. Source can be either quantitative (e.g., sensor data, images, gene arrays, laboratory tests) or qualitative (e.g., free text, demographics).

Storage issues

The storage is achieved using virtualization in big data where it holds set of sensor information, video, imaging scan, etc. there are many company which store the big data in healthcare system like Electronic Health Records(EHRs). EHRs describe patient treatments and outcomes are rich but underused information. Whereas the traditional health data centers capture and store an enormous amount of structured data concerning a wide range of information including diagnostics, laboratory tests, medications.

III. BIOINFORMATICS AND ITS BENEFACTION

Bioinformatics can be defined as the application of computer technology to the management of biological information. It covers a study of the inherent genetic information molecular structure etc. In this study of inborn genetic disease computer is used to gather, store, analyze and integrate biological and genetic information which can be applied to gene-based drug discovery and development. The first bioinformatics tool was BLAST (Basic Alignment Search Tools) it is used to examine new DNA or Protein sequence. BLAST [7] [8] compare new sequence to all sequence in database to find the most similar known sequence.

Delve Bioinformatics data space

To delve in bioinformatics is in another word to search hidden information from the stored data. The healthcare information deals with the best use of information, through the help of technology, in ameliorating healthcare, public health. When the data are delved correctly it gives the information about a patient’s state of health to right person, at the right time it is required, thereby guiding them in making and taking informed decisions about patient’s treatment.

Medical informatics makes available the essential tools to relate data and knowledge in the process of decision making. Information management and technology ensures collection from various sources, as well as processing and delivery of the same. The main aim of this delve data space in bioinformatics it increases the performance of treatment because the information has interpreted correctly.

IV. HEALTHCARE ANALYTICS

Analytics is the systematic use of data and related business insights developed through applied analytical disciplines such as statistical, contextual, quantitative, predictive, cognitive and other including emerging model to drive fact-based decision making for planning, management, measurement and learning. It may be descriptive, predictive or prescriptive. The big data guarantees the support of a large number of medical and healthcare which needs to be analyzed so they can support
disease surveillance and population health management [9]. It is also used to obtain various data from patient, patient being able to access their own data and personal health records and finally being able to work across these multiple data to identify patient and population health needs.

The analytics [10] in healthcare came as a result of large healthcare data that are being gathered electronically. Therefore, any technology or system is acclimated to improve the healthcare system. The recent human Genome Project [7] has found that the opportunities came to be able to develop such data, analyze, and present the genomic data in a quicker, cheaper and reliable way. This grand technological advance has affected healthcare in condition of new prevention, development, diagnosis, and ameliorated treatment for regular healthcare. This huge amount of data collected from bioinformatics and healthcare informatics [11] fields paired with analytics are appraised to save in near future preventive, predictive and personalized healthcare aids. For this high appraisal of analytics of big data in healthcare give tremendous aid to national prosperity and development of rural population. The best use of healthcare analytics [12] will provide the achievement of best output result. Analytics will provide therefore ameliorate efficiency, aid discovery and exploration of new drugs and advanced treatment.

V. CHALLENGES IN CLOUD SPACE FOR HEALTH CARE SECTION

Healthcare resources such as hospitals, medical practices and facilities need to be carefully considering the type of application clinical / nonclinical [13] when moving to cloud space. Security and confidentiality, consistency, amalgamation and data portability are some of the significant challenges and barriers to implement healthcare in cloud space but Cloud computing [24] with efficient Big data analytics [15] may end up being one of the most important IT advances to healthcare industry as a large and growing percentage of hospital management are storing data, clinical applications and email in cloud. This paper is expected to provide new and improved client care capabilities simultaneously limiting the increasing healthcare cost.

To prevent revelation, negotiation and abuse of the healthcare data maintained in the cloud space which may be personal, private or confidential requires proper protection.

VI. METHODOLOGY

The healthcare informatics focuses on health data, information and knowledge, including their collection, processing, analysis and use. Bioinformatics employ computational tools and techniques to study and analyze large biological databases and to absolutely to understand disease and study of inherent genetic information molecular structure by relating them with healthcare data [16]. This gathering, storing, analyze, and integrate biological and genetic information it leads to gene-based drug discovery and development. This paper touches on big data in healthcare, analyzes of those big data in healthcare for the better improvement healthcare system, bioinformatics data storage in secured manner.

This paper is a start up for a digitized health care system to move towards a cloud based technology [17] from current app-based solutions which include its own constraints and download costs. It also creates grounds for merging data analytics with pharma tie-ups leveraging the quality of EMR’s, currently which is a challenging task. This paper will look on the helpful result, the beneficence of each of them in amelioration of healthcare system.

The tactical recommendations [18] given in figure 3 is used to develop a valued and efficient HealthCare System incorporating the advantages of cloud computing are:

- **Understanding the value schemes** (Quantifiable benefits that cloud is delivering) for cloud operation-To improve the health care system the cloud computing technology is to be appendage with new Healthcare process and advances such as HIE (Health Information Exchanges), EMR (Electronic Medical Records), VCS (Virtual Care Solutions) and CHA (Connected Health Care Application) that can be used for patient care and patient outcome.

**Value Schemes can be:**

a. A **Healthcare infrastructure** that requires a significant upgrade to capture, share and protect EMR infrastructure which can be provided by cloud base offerings i.e. portable and open medical image archiving services from significant cloud providers enabling offloading of these tasks from hospital IT departments.

b. More **sophisticated Healthcare application** involving improved broad band capabilities should rely on cloud computing framework to perform the key back end related services, since mobile devices have limited computational capacity and storage capabilities.

c. **Monitored patient data** is maintained in cloud and can be tied to cloud base EMR data, these form of Remote Monitoring involving health related activities such as counting calories, calculating body mass index, exercise management, chronic disease monitoring, heart rate monitoring and smoke cessation can deliver new services to previously untapped markets as well as to provide new sources to the health care system.

d. **Big Data** which provides the mechanism for aggregating, mining, and analyzing large amounts of medical data on processes, treatments, effectiveness, costs and conditions. These cloud based approach can also be shared among authorized healthcare providers to provide access to potentially life-saving information.

![Figure 3 Tactical recommendations to develop a HealthCare System](image-url)
• **Surveying cloud workload and operation models**

Health care system operation implementing EMR initiatives in health system involves complexity. The new proposed infrastructure technologies that are facilitating cloud computing paradigm can reduce costs, improve performance and accelerate efforts to achieve meaningful use; the cloud based environment will lower the barrier for innovation and modernization of HealthCare systems.

Different cloud operation models offer different benefits, capabilities and risks that should factor into the development of Healthcares system. Despite the promising capability of cloud there are numerous technical approaches and factors (e.g. security, service level, availability, maintenance, system utilization) that must be considered. The public cloud [19] is suitable for health care system whereas the private cloud might be better suited for commodity enterprise workloads and applications like email, collaboration, financial system and resource planning.

• **Developing a cloud solution based on the value scheme, workload and operation model**

The healthcare system can share common infrastructure which reduces the cost of infrastructure due to resource sharing. Provisioning new resources, workloads or applications now becomes a faster process and can be controlled through a single IT administered control system. In health care system procurement process, infrastructure sharing multiple health care organizations naturally leads to standardization of the surrounding IT processes required to manage the infrastructure. Once standardization is achieved additional automation can be used to further shorten process times. A side benefit of automation is the reduction of human error, thus increasing service availability.

• **Developing a plan to supervise key performance indicators (KPI’s) to validate the payback of cloud based health care system**

The key performance indicators are generally classified as cost, infrastructure utilization, security, quality, availability, user satisfaction, audit controls, profitability indicators as it relates to cloud computing[25] characteristics and service delivery. Key performance indicators such as company specific metrics with traditional IT or potentially non-IT methods (e.g. paper based EMR, current imagery scan...
to be compared with the adopted cloud computing solution.

KPI’s support the overall healthcare related goals (e.g. patient care quality) will ultimately validate the role of cloud based solutions. The proposed model of the health care system is shown in figure 4 and also the Health Care system must ensure the following guidelines when transitioning to cloud computing as shown in figure 5.

- Health Care system must ensure the following when transitioning to cloud computing

Figure 5 Guidelines for Transition of HealthCare System to Cloud Space

a. Health Care system must be adaptable to various health care departmental and organizational sizes. Different healthcare organizations and departments will have their respective diverse sensitivities as to the uptime availability, system responsiveness, latency and scalability requirements for their healthcare IT applications and workflow.

b. Health Care architectures must encourage more open sharing of information and data sources. Standardization to be done in medical data formats and vocabularies, the adoption of secure mechanisms for widespread, interoperable information exchange between all of the healthcare players [21] such as providers, patients, government agencies, insurance companies.

c. Technology refreshes cannot overburden the already brittle budgetary environments. The healthcare organizations should first develop a strategy that complements its business goals / timelines with its current IT infrastructure and technology refresh cycle with the necessary elements of the cloud.

d. Scalability is a must as more patients enter the system and more data becomes digitized.

e. Portability is needed as doctors and patients would be benefiting from the ability to remotely access systems and data.

f. Security and data protection are paramount. Security and data privacy issues are of vital importance in the adoption of any IT-based healthcare solution. Appropriate security should be implemented and potentially enhanced not only in the existing IT infrastructure but into the underlying cloud monitoring and management process.

To achieve this health amelioration in bioinformatics we can use IoT (Internet of Things) to collect data from smart devices such as mobile, smart phone and intelligent medical devices; which is internetworking of physical devices and other items embedded with electronics, software, sensors, actuators and network connectivity which enables these objects to collect and exchange data. IoT is expected to offer advanced connectivity [28] of devices, systems and services that goes beyond machine to machine communications and covers a variety of protocols, domains and applications. We also consider Hadoop as a tool which collect and analyze the huge amount of data in healthcare system. This paper concludes with Hadoop as a health care model that provides public health care sector with specific guidance on how best to achieve the benefits of cloud computing with efficient big data analytics while maintaining permissible level of risks.

This paper suggests proposed model with various components for health care system:-

Component 1: Rating a doctor efficiently and rapidly whose details are limited to medical sorority alone.

Component 2: Patient and hospital management which is mainly focused on patient’s data to run big data analytics on the same assisting doctors to streamline their booking, payments etc.

Component 3: A peer-to-peer experiential learning facility where large number of doctors can interact per day discussing their cases and learning more about the advancements through peer-to-peer knowledge sharing.

Component 4: New encryption technologies that protect resources required to meet stringent regulatory requirements.

Component 5: The component used to educate people for cheaper alternative for same generic drugs prescribed by the doctors.

Some of the objectives of above attributes can be achieved using advanced neural network algorithms [26], data mining algorithms, heuristic search algorithms for classification of health care data’s. This ensures the importance of computational methods [6] [14] [20] such as Neural-Fuzzy mapping in bioinformatics and computational biology that are much cheaper and faster than conventional experimental methods.

VII. OPEN SOURCE CLOUD FRAMEWORKS

The healthcare informatics focuses on health data, An Open Source Java Framework technology called Hadoop (Highly Archived Distributed Object Oriented Programing) was created by Goug Cutting and Mike Cafarella in 2005 for supporting a distributed search Engine Project. It helps to store, access and gain large resources from big data in a distributed fashion at less cost, high degree of fault tolerance and high scalability. Hadoop [22] handle large number of data from different system like images, video, folders, sensor records, Communication Data, and anything which we can’t think in any format.

All these resources can be stored in Hadoop cluster without any schema representation instead of collecting from different system. Hadoop contains one master node and many slave nodes. The master node consists of data node which holds node, Job Tracker and Task Tracker where are slave node acts as Task Tracker and Data Node which holds compute only and data only worker node. The job scheduling is managed by the Job Tracker.
Hadoop Distributed Files system (HDFS) and Map Reduce are the two major parts of Hadoop. Hadoop Distributed Files system provides storage of data. It is more scalable, reliable, distributed in Hadoop framework environment. It provides the operation of “Write Once, Read Many Times”. The files are stored as block with default size of 64MB. The communication between nodes occurs through Remote Procedure Calls.

- **Security concerns with Cloud Frame works**

Hadoop was primarily deployed on-premise very shortly in most of the enterprises. Security is not an issue when sensitive data was safely confined in isolated clusters or data silos. But things have been changed as Hadoop evolved into Big Data as -a - Service (BDaaS) and took it to the cloud, and became bounded by a constantly growing ecosystem of tools and applications. When innovations have provided to publicize the data and bring Hadoop into the mainstream, they have also created new security issues for organizations that now fight to scale security in step with Hadoop’s rapid technological progresses.

Hadoop has evolved into an enterprise data platform for many organizations. As data that was once stored is brought together in a vast data lake and made accessible to a variety of users across the organization it gives rise to new security challenges. Among these challenges are:

- Make sure the proper validation of users has been done who access Hadoop.
- Make certain that authoritative Hadoop users can only access the data that they are permitted to access.
- Guarantee that data access histories for all users are recorded in accordance with compliance regulations and for other important purposes.
- Assure the protection of both types of data i.e. data both at rest and in transit with the help of enterprise-grade encryption.

VIII. CONCLUSION

It has been shown that biology and technology in the term of bioinformatics is to provide the opportunity for back and forth in technology, data, and knowledge. This will help in determining the database of diseases, their causes, the opened way of treating those disease and cure which will prevent the spread of those maladies. Security and confidentiality, consistency, amalgamation and data portability are some of the significant challenges and barriers to implement healthcare in cloud space but Cloud computing with efficient Big data analytics may end up being one of the most important IT advances to healthcare industry as a large and growing percentage of hospital management are storing data, clinical applications and email in cloud. By using Hadoop framework, there are other possible cloud based platforms [27] which can be used for Health Care System (HCS) such as Apache Flume, MillWheel and Google’s won Cloud Dataflow. Also these platforms have in common the ability to improve the efficiency and reliability of data collection, aggregation and integration. The analytical point of view in bioinformatics will be improved and the life of nation will be ameliorated through the reduction of disease, identify future diagnostics.

IX REFERENCES

[1] AmirGandomi, MurtazaHairara, "Beyond the hype: Big data concepts, methods, and analytics", International Journal of Information Management, vol 35, Issue 2, pp. 137-144,2015.
[2] Javier Andreu-Perez, Carmen C.Y. Poon, “Big data for Health”, IEEE journal of biomedical and health Informatics, vol. 19, No. 4, pp.1193-1208, 2015.
[3] Philip Ball, "Counting Google searches predict market movements", www.nature.com / news / counting-google-searches-predicts-market-movements-1.12879, Nature Journal, doi:10.1038/nature.2013.12879, 2013.
[4] Christopher Matthews, "'Trouble with Your Investment Portfolio? Google's Io"", TIME Magazine, www.lik2do.com / learn?is=Big_Data, 2013
[5] Bernhard Warner, "'Big Data' Researchers Turn to Google to Beat the Markets", www.bloomberg.com / business week, Bloomberg Business week, 2013., 2013.
[6] Mukunthan. B and Pushpalatha. A., “Automation of DNA Finger Printing for Precise Pattern Identification using Neural-Fuzzy Mapping Approach”, International Journal of Computer Applications, vol. 13(3), pp.16-24 2011.
[7] Kmiecik, Sebastian; Gront, Dominik; Kolinski, Michal; Wieteska, a; Kasz; Dawid, Aleksandra Elzbieta; Kolinski, Andrzej, “Coarse-Grained Protein Models and Their Applications”, Chemical Reviews, vol. 116(14), pp. 7898-7936, doi:10.1021/acs.chemrev.6b00163,2016.
[8] Joyce, Adam P.; Zhang, Chi; Bradley, Philip; Havranek, James J, "Structure-based modeling of protein: DNA specificity", Briefings in Functional Genomics, Vol. 14 (1), pp. 39–49, doi: 10.1093 / bfgp / elu044,2015.
[9] David Leinweber , "Big Data Gets Bigger: Now Google Trends Can Predict The Market", Forbes, 2013.
[10] HiraKashyap, Hasin Aizal Ahmad "Big Data Analytics in Bioinformatics: A Machine Learning Perspective", Journal of latex Class Files, vol. 13(9), pp. 1-20, 2014.
[11] JeliliOyelade, JunokeSoyeni, "Bioinformatics, Healthcare Informatics and Analytics: An imperative for Improvement Healthcare System", International Journal of Applied Information System, Foundation of Computer Science FCS, New York, USA, vol. 8(5), pp. 1-6, 2015.
[12] Arushi Jain, Vishal Bhatnagar “Crime Data Analysis Using Pig with Hadoop", International Conference on information Security & Privacy(ICSIP2015), vol. 78, pp. 571-578, 2016.
[13] Cloud Standards Customer Council: Security for Cloud Computing Ten Steps to Ensure Success, www.cloud-council.org / resource-hub, pp. 1-35, 2015.
[14] Mukunthan. B and Nagaveni. N, “Automating Identification of Unique Patterns, Mutation in Human DNA using Artificial Intelligence Technique”, International Journal of Computer Applications, vol. 25(2), pp. 26-34, 2011.
[15] “HADOOP Big Data Analysis Framework”, www.tutorialspoint.com / hadoop / hadoop tutorial.pdf, 2014.
[16] Richard Waters, "Google search proves to be new word in stock market prediction". Financial Times, 2013.
[17] Brian T.Horowitz, "Cloud Computing in Health Care to reach $5.4 Billion by 2017", www.eweek.com, -Report- 512295, 2012.
[18] Silvia Piai, “Cloud in the Western European HealthCare Sector: Trends and Strategies for 2012 and Beyond”, IDC Health Insights, 2012.
[19] Markets and markets.com- Healthcare Cloud Computing (Clinical, EMR, SaaS, Private, Public, Hybrid) Market-Global Trends & challenges, Opportunities & Forecasts , www . marketsandmarkets.com /Markets-Reports/cloud-computing - healthcare-market-347.html, 2017.
[20] Mukunthan. B and Nagaveni. N, “Identification of Unique Repeated Patterns, Location of Mutation in DNA Finger
Printing Using Artificial Intelligence Technique”, International Journal of Bioinformatics Research and Applications, vol. 10(2), pp. 157-169, 2014.

[21] Reichman, O.J.; Jones, M.B.; Schildhauer, M.P. (2011). "Challenges and Opportunities of Open Data in Ecology". Science, 331 (6018): 703–5. doi:10.1126/science.1197962. PMID 21311007.

[22] B. Saraladevi, N. Pazhaniraja, P. Victer Paul, “Big Data and Hadoop-A Study in Security Perspective”, 2nd International Symposium on Big Data and Cloud computing (ISBCC’15), vol. 50, pp. 596-601, 2015.

[23] Hamish McRae, "Hamish McRae: Need a valuable handle on investor sentiment? Google it", The Independent, London, 2013.

[24] Sanjay P. Ahuja1, Sindhu Mani1 & Jesus Zambrano, “A Survey of the State of Cloud Computing in Healthcare”, Network and Communication Technologies, vol. 1(2), pp.12-19, doi:10.5539/ntc.v1n2p12, 2012.

[25] The Power of cloud, “Driving Business model innovation”, www.ict-industry-reports.com / wp-content / uploads / 2012 / 03 / 2012- IBM-cloud-report - March2012.pdf, pp.1-19, 2012.

[26] Mukunthan, B, “A Neural Network Approach for Precise Pattern Identification of Human DNA”, International Journal of Neural Networks and Applications, International Science Press, vol. 3(2), pp. 55-62, 2010.

[27] Big Data Problems, http://adtmag.com/whitepapers/2015/12/ibm-security-guardium-securing-big-data-pursuit-big-opportunities.aspx.

[28] Krishnakumar K.G, Dr.B.Mukunthan, “Cross Layer Based Adaptive Routing Approach for VANET”, International Journal of Control Theory and Applications, vol .9, Issue 40, pp. 161-169, 2016.