ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 07, Issue, 07, pp.14204-14207, July, 2017

ORIGINAL RESEARCH ARTICLE Open Access

AN APPROACH TO PROVIDE SECURITY AND PRIVACY IN E-HEALTH CARE SYSTEM USING HADOOP MR

***[1]Prasanna, B. and [2]Prema, A.**

[1]Research Scholar, Department of Computer Science, Sivaganga, Tamilnadu, India
[2]Assistant Professor, RDM Government Arts College, Sivaganga, Tamilnadu, India

ARTICLE INFO

Article History:
Received 15th April, 2017
Received in revised form
07th May, 2017
Accepted 26th June, 2017
Published online 31st July, 2017

Keywords:
E-Health care, Big Data,
Cloud, Hadoop, Map Reduce.

****Corresponding author: Prasanna, B.,***

ABSTRACT

Big data and cloud computing is the backbone to the modern Data Storage, which provide a perfect security for data repository in the recent systems and managed with the cloud computing. Big data is a term for data sets that are so large or complex that traditional data processing application software's are inadequate to deal with them. Challenges include capture, storage, analysis, data curation, search, sharing, transfer, visualization, querying, and updating and information privacy. Hadoop is the core platform for structuring Big Data, and solves the problem of making it useful for analytics purposes. It is designed to scale up from a single server to thousands of machine, with a very high degree of defect easily. In this paper, we propose an e-health protection framework using Map Reduce Technique.

Citation: Prasanna, B. and Prema, A. 2017. "An approach to provide security and privacy in e-health care system using hadoop mr", *International Journal of Development Research*, 7, (07), 14204-14207

INTRODUCTION

The main functions of the cloud server are to create interface between application and user. The user username and password then only they can able to connect the server. If the user already exits directly can login into the server or user must register their details such as username, password, and Email id, keen on the server. The server will create the account for the entire user to maintain upload and download rate. The term "big data" often refers to the use of predictive analytics, user behaviour analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. "There is little doubt that the quantities of data now available are indeed large, but that's not the most relevant characteristic of this new data ecosystem."
Hadoop consists of a storage part, known as Hadoop Distributed File System (HDFS), and a processing part which is a Map Reduce programming model. Hadoop splits files into large blocks and distributes them across nodes in a cluster. It then transfers packaged code into nodes to process the data in parallel. This approach takes advantage of data locality – nodes manipulating the data they have access to allow the dataset to be processed faster and more efficiently than it would be in a more conventional supercomputer architecture that relies on a parallel file system where computation and data are distributed via high-speed networking. A Map Reduce program is composed of a Map procedure (method) that performs filtering and sorting (such as sorting students by first name into queues, one queue for each name) and a Reduce method that performs a summary operation (such as counting the number of students in each queue, yielding name frequencies). The "Map Reduce System" (also called "infrastructure" or "framework") orchestrates the processing by marshalling the distributed servers, running the various tasks in parallel, managing all communications and data transfers between the various parts of the system, and providing for redundancy and fault tolerance. Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services), which can be rapidly provisioned and released with minimal management effort.

Section 2 of this paper deals with related work done in e-health care electronic health record, google map reduce,and hadoop distributed file system. Section 3, explains an actual process of e-health care, and proposed work, in section 4, experimental analysis and implementation results are given, and finally, section 5 presents a conclusion of this paper.

Related Work

Zhang Xin, et al explained, Cost Effective, on-demand self - service, Ubiquitous network access, fast flexibility, elevated dependability, flexibility in cloud computing performance, accessibility and safety are major research topics. Among them, cloud computing security is one of the important research topics (Zhang Xin, 2012). Han Hu, et al pointed out that the Map Reduce is useful in a number of areas where huge data analysis is required (Han Hu, 2014). Tom, White clarified many applications which handle the huge data sets have been developed (Tom White, 2012). R.Agrawal, et al described that, E-health has increased its success and popularity in a short period of time. In practical, the system has to be secured and e-health service provider is entrusted with the responsibility to handle the sensitive information (Agarwal, 2004). D. C. Kaelber et al stated, now organizations invest millions of dollars in the best PHR (Personal Health Record) architectures, value propositions, and descriptions (Kaelber, 2008). Ming, et al presented the personal health record is used to store the data onto the user in secure and in an efficient way. It will be a valuable asset to individuals and family, enabling them to combine and manage their healthcare information using safe, standardized tools (Ming and Wenjing Li, 2012). Individuals own and manage the information about the PHR, which comes from healthcare providers. The personal health record has been maintained by a private environment so that only authorized user can access the data. The PHR does not replace the legal record of any provider (Alexander Dobrev, 2009). D. T. Mon, et al described in others words; they are electronic health records (EHRs) that are owned by patients,which contain medical data generated within one specific care institution (Health Informatics- Electronic Health Record- Definition, 2005 and Mon, 2013). Haluk Demirkan illustrate Electronic Health Records Systems (HER) store the entire patient's medical history information about the time of an admission, medical tests performed on the patient, medicine prescriptions, readmission information and any other relevant information of the patient. These data has to be availabled and managed by all health care providers (Haluk Demirkan, 2013). Nowadays, the data that need to be analyzed are not just large, but they are collected of various data types, and even including streaming data (Russom, 2011). Although it seems that big data makes it possible for us to collect more data to find more useful information, the truth is that extra data do not automatically mean more useful information. It may contains more unclear or irregular data. For instance, a user may have various accounts, or an account may be used by various users, which may degrade the correctness of the mining results (Boyd, 2012). Therefore some new issues for data analytics arrive up, such as privacy, safety, storage, fault charity, and worth of data (Katal, 2013). Map Reduce is a framework of efficiently processing the analysis of big on a large number of servers. It was developed for the support for Google's search engine to enable number of commodity servers to efficiently process the analysis of huge numbers of web pages collected from all over the world (Dr. Siddaraju, 2014). Apache (http:hive.apache.org, and http:giraph.apache.org) developed a project to implement Map Reduce, which was published as open source software (OSS), this enables lots of organization, such as business and university, to tackle big data analysis .Map Reduce (Guoping Wang and Chee Yong Chan), is a simple programming model for processing he data sets in parallel. The GMR (Google Map Reduce) was invented by Google back in their earlier days so they could usefully index all the rich textural and structural information they were collect, and then present significant and actionable results from users. Map Reduce (you map the operation out to all of those servers and then you reduce the results back into a single result set), is a software paradigm for dispensation a large data set in a distributed similar way (Prashant Chauhan Abdul Jhummarwala, 2012). A small Hadoop cluster includes an only master and several employee nodes. The master node consists of a Job follower, Name node and Data node. A slave or employee node acts as both a Data nodule and Task follower, though it is possible to have data-only employee nodes and compute-only employee nodes. These are usually used only in unusual applications. Hadoop requires safe Shell to be between nodes in the cluster. In a big cluster, the HDFS is managed through a committed Name Node server to host the file system index, and a secondary Name Node that can create snapshots of the name node's memory structures, thus preventing file-system fraud and tumbling failure of data (Silky Kalra, 2014).

Proposed Work

The proposed e-health care framework is presented in Figure 1. In this framework, firstly, the user logs on by using username, password and by providing unique biometric information to secure user identity. Secondly, Request for Service is generated by the admission department to get request for the database and external third parties. Thirdly, after administrator requests for the service, health requester who asks for access to store personal health information is known as Data can decrypt the encrypted Personal Health Information (PHI). Fourthly, after requesting to access the PHI, then Single point of contact checks whether the user is allowed to access the service or not. If the user gets allowed to access the service, administrator can easily access the patient coming from external hospital or referred by doctor. Then Personal Health Information is transferred to Health Service by using a new proposed framework MA-ABE (Multi Authority Attribute based Encryption). In our current framework, the encryption technique CP-ABE(Cipher Policy Attribute Based Encryption)that was used to secure e-health system is average as new enhanced encryption technique can be used to overcome this technique which provides find access control with excellent as compared to CP-ABE. In addition, its efficiency is scalable as compared to current solution and it's collision resistant.

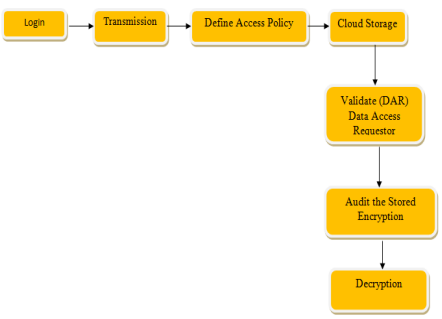

Figure 1. Methodology for proposed E-Health Care

Figure 2 describes the patient and Health service provider shared the key having public and private key. Patient sends encrypted information about image of the health service provider. After receiving the message from the patient, Health service provider decrypted the message and ensure by validating the image, if it is verified then health service to provide to encrypt personal health information according to patient defined policy. After the transmission of personal health information about health service providers, PHI is stored at cloud. To get access to sensitive information data access requester sends access request to cloud service providers and then health service providers executes the access request. The current selected solution has changed features. It ensures patient identity privacy, make sure PHI integrity and source authentication, simply revoke malicious user, resistant to collusion attack and, patient-centric access control.

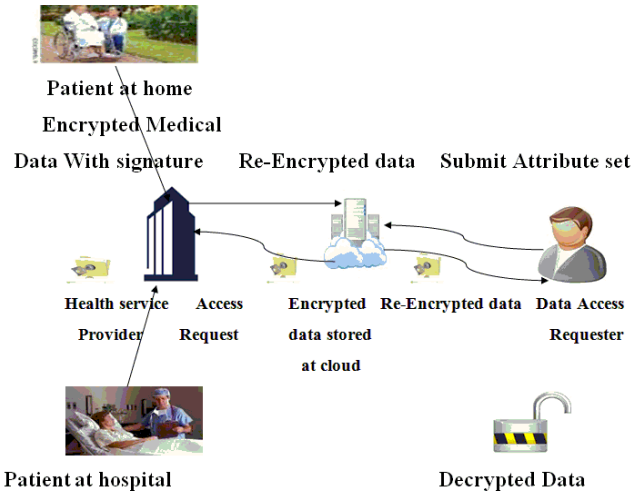

Figure 2. Proposed frameworks for E-Health Environment

Input Table: The input data are given in the Table 1.

Table 1. Input Table

Field	Type	Null
Name	varchar(100)	No
Age	varchar(100)	No
Address	varchar(100)	No
State	varchar(100)	No
Country	varchar(100)	No
Username	varchar(100)	No
Password	varchar(100)	No
Mail id	varchar(100)	No
Photo	varchar(1000)	No
Category	varchar(1000)	No
ID	varchar(10000)	No

This approach has some limitation. Collision resistant cannot completely be important to this encryption method, and Fine gained access control is average in the CP-ABE method. Good organization in this framework is not scalable, and Average authentication and approval is limitation.

Implementation Result

This Proposed method can be used by any organisation to store their patient medical information. PHR allows the controlled sharing of application software that is required to view and analyze health records. Here the patient data is shared securely among cloud using ABE (Attribute Based Encryption) technique. Using this technique each user generate one private key depends upon the user data. The security key is used to retrieve the uploaded data. Suppose one user uploads many data means the same authorized key is used to retrieve that user data. All that keys are securely maintained by individual. This key will be shared by particular user via e-mail id.

The following picture shows the urban and rural wise heart diseases analysis.

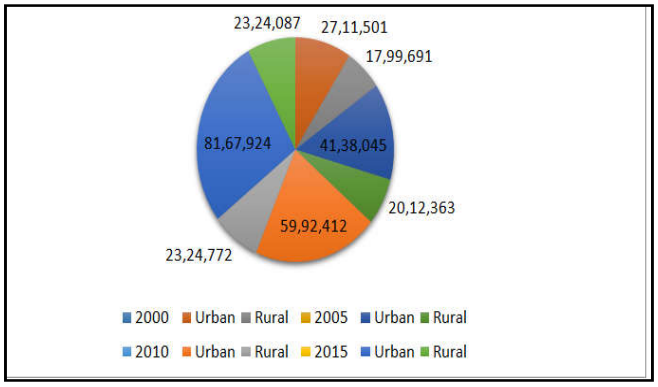

Figure 3. Heart Diseases analysis

By using to attribute based encryption, we reduce the key management problem and also enhance privacy guarantees which play an important role because these are unique and not easily hackable.

The following picture shows the state wise Tobacco analysis.

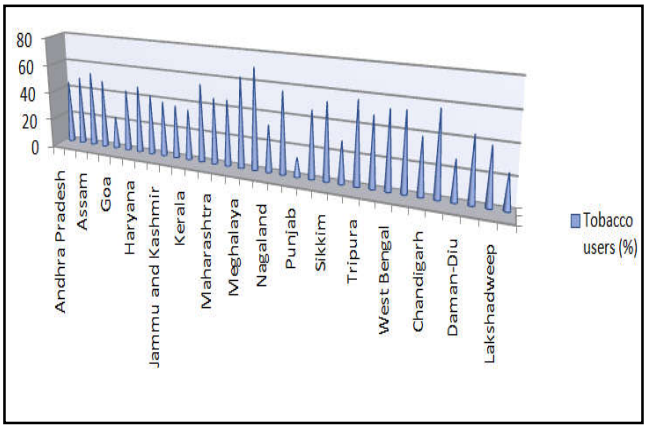

Figure 4. Tobacco analysis

The following picture shows analysis of diseases.

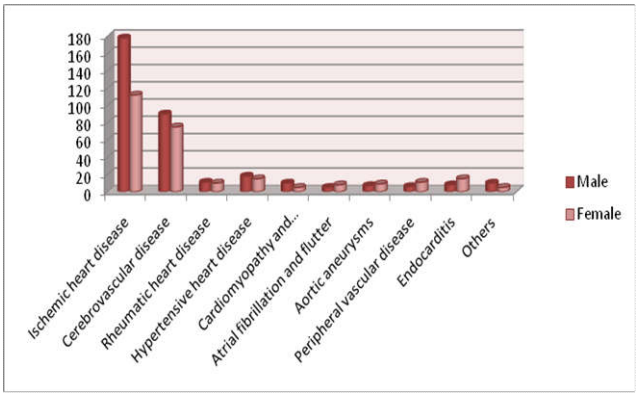

Figure 4. Diseases analysis

Output Table: The Result are shown in Table 2.

Table 2.

Name	Age	Address	State	Country	Username	Password	Mail id	Photo	Category	Id
V	87	Madurai	Tamil Nadu	India	V	V	v@gmail.com	Photo-3.jpeg	Patient	40
Kalai	67	Madurai	Tamil Nadu	India	Kalai	kalai	kalai@gmail.com	Photo-4.jpg	Patient	7
Nagu	25	sivagangai	Tamil Nadu	India	Nagu Prasi	naguprasi246	naguprasi6@gmail.com	images.jpg	Hospital persons	45
Ravi	56	Chennai	Sikkim	India	Ravi	ravi	ravi@gmail.com	Photo-5.jpg	Doctor	18
John	47	Mumbai	Nagaland	India	John	john	john@gmail.com	Photo-11.jpg	Doctor	1
Joe	67	Mumbai	Rajasthan	India	Joe	Joe	joe@gmail.com	Photo-13.jpg	Patient	6
Johnson	67	Bangalore	Tamil Nadu	India	Johnson	Johnson	johnson@gmail.com	Photo-1.jpg	Patient	16
Larance	67	Bangalore	Tamil Nadu	India	Larance	Larance	larance@gmail.com	Photo-2.jpg	Patient	32

Conclusion

Security in health care should increase quality of health care and reduce the cost of health care. Our research focuses on securing health care from attacks by unauthorized users and also identifies the threats and vulnerabilities in health care. It highlights the approach and technique that is equally used in security Health IT and also reflect cyber security threats. This proposed research describe in securing health records of the efficient and scalable by using Attribute in Based Encryption (ABE) techniques for securing PHR.

REFERENCE

"Health Informatics- Electronic Health Record- Definition, Scope and Context*", International Standards Organization,* iso/tr 20514:2005, Jan 2005.

Agarwal, R., A. Kini, K. LeFevre, A. Wang, Y. Xu, D. Zhou, "Managing Healthcare Data Hippocratically", Proc.Of Acm Sigmod International Conference on Management of Data, Paris, France, June 2004.

Alexander Dobrev, Tom Jones, Karl Stroetmann, Yvonne Vatter, Kai Peng empirica Communication & Technology Research, Germany; TanJent Consultancy, UK "The conomic impact of interoperable electronic health records and e prescription Europe ", Dg Infso &Media 30-CE-0161851/00-30 October 2009

Apache Giraph Project, http:giraph.apache.org

Appache Hive, http:hive.apache.org

Boyd d, Crawford K. "Critical questions for big data". Inform Communication Soc.2012:15(5):662-79

Dr. Siddaraju, Sowmya C L, Rashmi K, Rahul M"Efficient Analysis of Big Data Using Map Reduce Framework" International Journal of Recent Development in Engineering and Technology Website: www.ijrdet.com (issn 2347-6435(Online) volume 2, Issue 6, June 2014).

Guoping Wang and Chee Yong Chan, MultiQuery Optimization in Map Reduce Framwork.{wangguoping, chancy}@comp.nus.edu.sg

Haluk Demirkan , "A Smart Health care system Framework", Ieee, oct 2013.

Han Hu, Yonggang Wen, Tat-Seng Chua, Xuelong Li "Towards Scalable system for big data analytics", june 24 2014 Digital Object Identifier 10.1109/ACCESS.2014.2332453 ieee.

Kaelber, D.C., A.K. Jha, D.Johnston, B.Middleton, and D.W. Bates,- Viewpointpaper: "A research agenda for personal health records(phrs)", *J.Amer. Med Inform. Assoc.,* Vol 15, no. 6, pp 729-736, 2008.

Katal A, Wazid M, Goudar R. Big data:issues,challenges,tools and practise. In:Proceeding of the International Confrence on Contemporary Computing, 2013. Pg 404-409.

Ming and Wenjing Li, Shucheng Yu,Yao Zheng, Kuiren, Lou, -"Scalable and Secure Sharing of Personal Health Record in Cloud Computing using Attribute-based Encryption "ieee transactions on parallel and distributed systems vol. xx, no. xx, xx 2012

Mon, D.T., J.Ritter, C.Spears, and P.Van Dyke,-HL7 PHR system functional moel, May 2008

Prashant Chauhan Abdul Jhummarwala, Manoj Pandya,- "Detection of ddos Attack in Semantic" Web International Journal of Applied Information System(ijais)-issn: 2249-0868 Foundation of Computer Science fcs, New York, Usa Volume 4-No.6,December 2012-www.ijais.org.pp.7-10

Russom, P,"Big data analytics". Tdwi 2011 Vol. 3 No. 6, 137-142.

Silky Kalra1, Anil lamba2 "A Review on Hadoop Map Reduce- A Job Aware Scheduling Technology" issn (e):2250-3005 vol. 04 Issue ,5 May – 2014 *International Journal of Computational Engineering Research* (ijcer)

Tom White, Hadoop: The Definitive Guide, O' Reily Press 2012-01-27 vol isbn**:** 978-1-449-31152-0 1327616795

Zhang Xin, Lai Song-Qing, Lai Nai-wen." Research on Cloud Computing Data Security Model Based on Multidimension" 978-1-4673-2018-2112@2012 ieee

****