Medical (Healthcare) Big Data Security and Privacy Issues

Saurabh Pandey, Rashmi Pandey

Abstract— Big data is one the most promising, essential and modern computing era around the world. Big data is most commonly described as huge amount of un-structured data or we can say semi-structured data. The processing, storage, and analysis of such huge sets of data with the help of classical processing or database approaches or tools are insufficient, it requires advanced processing tools with real-time analysis. Healthcare (medical) is one of the important sector that produces big data because today, healthcare switches paper-based medical records into electronic platform to store, manage, analysis and process in the form of Electronic Medical Record (EMR) or Electronic Healthcare Record (EHR) with the help of internet..Due to vast digitization of medical big data(records) or we can say the use of technologies to transform healthcare industry into electronically available data to consumers, there are various security and privacy risks associated with the use of such technologies must be addressed in order to provide better and safe access of records online and make the system acceptable worldwide with the contribution of healthy economic growth. In this paper, we explore various security and privacy concerns surrounding medical (healthcare) big data.

Index Terms— Big data analytics, Electronic Patient Record (EPR), Electronic Health Record (EHR), Electronic Medical Record (EMR), Health Information Technology (HIT), Magnetic Resonance Imaging (MRI), Security and Privacy

1 INTRODUCTION

AS we know, the new generations of data are most widely available in the form of electronic mode that makes a way of digitizing healthcare (medical) records. Due to large frequency of generating such kind of big data or records, the healthcare industry is witnessing many challenges in terms of complexity of handling as well as processing such big data. Also with the rapid increasing costs of medical (healthcare) systems and healthcare medical insurance premiums there is a need to switch from conventional systems in the form of computation, storage and communication system to novel advanced computation, storage and communication systems. As we know, cost reduction is must in order to enhance efficient healthcare delivery, process, analysis and management. Big data gives a reliable solution with the use of advanced technologies to transform current medical (healthcare) industry. McKinsey Global Institute estimates a $100 billion increase if big data analytics are used [1].Genomic research also enhance the physicians to make real time decisions and supervision on various treatments. Further, big data enables the development of predictive analysis models. Prediction is based on electronically available records i.e. clinical lab reports, MRI (Magnetic Resonance Imaging) and other factors as per the disease. Today, medical (healthcare) industry move from a volume-based industrial model to a value based industrial model, the value of data makes an important role in relation to patient care, and associated cost includes insurance claim cost [2]. Here, value based model can be described as-"To improve the quality of care and prediction analysis in order to improve decision making regarding treatment"[2]. In reference to the use of technologies paves the way in the field of Electronic Health Records (EHR) or Electronic Patient Record (EPR), Healthcare Information Technology (HIT). In medical systems, large amount of patient data (i.e. Social, Financial, Physical, Clinical, Environmental, Psychology and Genomics etc.) is essential in real time analysis, prediction, decision making in order to provide good patient care, all the data described above directly affect the condition of any patient that must be analyzed while performing further treatment process (Figure 1).

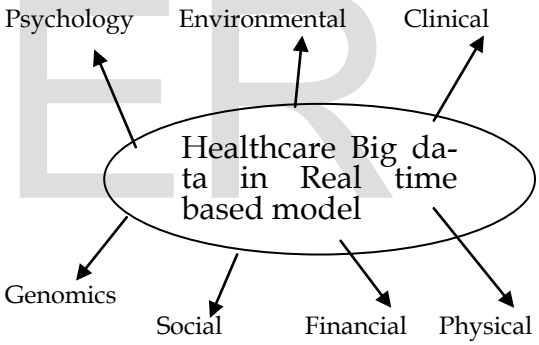

Figure1. Real time healthcare data [3]

Sensor network is also another technology adopted by medical industry. Here wearable sensors are used to examine patient at home with the help of remote patient monitoring system. For e.g. Intel's Integrated Hospital is aimed to improve healthcare in combination with mobile point-of-care (MPOC) and other information technologies to provide digital view of any patient condition [4].

Such technologies open the door of various security and privacy risks that must be explored in order to promote big data in medical (healthcare) industry.

In section 2, we describe security vs. privacy Section 3 discusses security and privacy issues in medical data in transmission with current system challenges. Section 4 discusses some existing solutions. The last section 5 concludes the study with some recommendations.

2 Security Vs. Privacy

Privacy in medical (healthcare) can be defined as-"The right and desire of any patient to control or regulate the disclosure of personal health information (Rindfleisch T.C.1997)" [5]. Privacy can also be defined by The Code of Fair Information (FIP) as-"There must be a path or way for any patient to prevent information about them that was obtained for one purpose from being used or made available for other purposes without the patient's consent." Patient consent is necessary in order to provide privacy. Consent includes who can access patient's particular record with valid purpose.

Security can be described as physical and technological measures that can be used to protect healthcare data from unauthorized disclosure or illegal access of any restricted data. Here access control policies are violated that must be prevented.

There are some security goals that can be pointed out as per ISO-OSI standard recommendation model are-Data Confidentiality, Data Integrity, Access Control, Authorization and Authentication.

These goals must be implemented in proper way to ensure data security in healthcare. To implement such security goals or policies provides privacy of patient data. Therefore, we can say that-Privacy deals with various security policies and security describes the tools or safeguard techniques to implement such security policies [5].

3 Security and Privacy Issues

Use of technologies such as: Electronic Patient Record (EPR) systems or we can say Electronic Health Record (EHR), Remote patient monitoring using sensor networks and the combination of electronic record with sensor networks as a hybrid can help to enhance the quality of medical process. There are many security and privacy concerns that must be analyzed inside this section:

3.1 Medical Big Data Storage and Access

As we know medical big data records are most widely in the form of EPR and remote patient monitoring using sensor networks, the records becomes vulnerable to hackers, malicious attacks and unauthorized access which increases the danger to hurt privacy and confidentiality[6]. There are two approaches to store data –centralized which store data in central database or it can be stored in local database connected to each database resides inside a network. Selection of storage schemes is critical to accommodator to handle security and privacy issues.

There are an emergency situations require disclosure of medical data without any consent by the patient to provide necessary care in minimum time limit [2]. The access policy needs to be evaluated according to that situation. To access any patient record in EPR, users are classified into two categories according to their privileges to access records:

Edit-Read/Write privileges (e.g. doctor, nurses etc.)

Read only privileges (e.g. medical insurance provider)

3.2 Lack of Public Trust

The medical data should be disclosed to right authorized person to enable the treatment. Healthcare entities limit the disclosure of record and implement proper procedures and policies to define appropriate level of access. Trust is necessary to increase the value of data. In Information Technology (I.T.), White more described two ways to provide trust as-agreements between participating entities involved in communication and the reliable, efficient and correct operations in I.T. solution [5].

3.3 Conflict between various security and privacy regulations/legislations/legal frameworks and standards

In current technological healthcare (medical) environment, there are many security and privacy regulations/ legislations/ legal frameworks and standards:

- HIPAA (The American Health Insurance Portability and Accountability Act) for security policies.[7][10] At the same time, there are many state legislations/ laws identified for e.g. Health Privacy Project, The state of health privacy and The Health Information Technology for Economic and Clinical Health Act (HITECH), which contain separate as well as conflicting laws and opens the door for security threat [8].
- There are a huge amount of international standards inherited by different geographical regions for e.g. HL7 (High Level International version 7), ICD (International classification of disease) [5] and many others which also increase the probability of malicious attacks.

3.4 Big Data Mining and Warehousing,Web Mining Technologies

To analyze and identify big data with the help of dependencies in the form of relationships and patterns is said to be big data mining and storage such patterns in data ware house. Web mining is integrated information originates with the help of classical mining methodologies. Mining healthcare big data requires rigid methods and constraints to store, manage and analysis in order to maintain security and privacy. Proper supervision and/or evaluation of mining techniques is a major challenge to increase efficiency.

4 Solutions

Here, let us we discuss some solutions discussed by researchers in concern to medical (healthcare) big data security and privacy issues:

4.1 Role Based Access Control/Security

One of the advanced essential model for access control. Access is the capability of any user to perform specified task. Roles are assigned according to their authorization of what amount of resources can be provided or granted.[9] Access control list (ACLs) must be refined time to time.

4.2 Use of Cryptographic Encryption Techniques

Encryption is the process of transforming plain text message to some form that can be readable only by communication entities involved in data transmission. Levels of encryption can be performed both software as well as hardware. There are various encryption algorithms such as Triple DES (Data Encryption Standard), AES (Advance Encryption Standard), IDEA (International Data Encryption Standard) etc.

4.3 Authentication

There are a number of authentication algorithms used for e.g. Digital signature, Password mechanism, in sensor networks hash functions can be used for authentication. The most popular algorithm can be used in sensor networks is Tinysec.[11] developed by Kensas State University.

4.4 Review Plan for Security and Privacy

Here security policies are applied based on organization requirement. Self control security needs to be applied.

5 Conclusion

As we know healthcare industry generates big data the security and privacy issues must be explored and prevented in order to transform medical (healthcare) industry. Some more works needs to be required for some areas such as Consent mechanism, Common legislation, Uniform medical big data storage format, Web and data mining technologies etc.
These factors needs some work in order to make big data medical industry worldwide acceptable.

References

[1] P. Groves, B kayyate,D knott and S.V. Kuiken,The big data revolution in healthcare, Mc Kinnsley & Company,2013

[2] Marci Meingast, Tanya Roosta, Shankar Sastri Security and Privacy Issues in Healthcare Information Technology, 2014

[3] Harsh Kupwade Patil and Ravi Seshadri, Nanthealth, Big Data Security and Privacy Issues in Healthcare,2016

[4] Solutions of Improving Healthcare at http//www.intel.com/business/bss/industry/healthcare/index.html

[5] Fatemeh Rezaeibagha, Privacy and Security of Eleectronic Patient Record Sharing, 2013

[6] Computer Science and Telecomminications Board, Networking Health: Perceptions for the internet, 2000

[7] " The American Health Insurance Portability and Accountability Act," U.S. Government Printing Office, 1996. [Online]. Available: http://www.gpo.gov/fdsys/pkg/PLAW- 104publ191/html/PLAW-104publ191.html.

[8] "Health Information Technology for Economic and Clinical Health Act," 2009. [Online]. Available:http://www.gpo.gov/fdsys/pkg/BILLS-111hr1enr/pdf/BILLS-111hr1enr.pdf.,

[9] Role Based Access Control at http://csrc.nist.gov/rbac

[10] HIPAA 101.com - Info Guide to HIPAA Compliance, Implementation and Privacy at www.hipaa-101.com

[11] C. Karlof, N. Sastry, and D. Wagner, TinySec: A Link Layer Security Architecture for Wireless Sensor Networks, Conference on Embedded Networked Sensor Systems, 2004.

Author Details :

- Saurabh Pandey is M Tech. Computer Science & Engg. Passout in 2017 from Integral University, Lucknow U.P. India. PH- 91-9889541421 E-mail: saurabhpdy16@gmail.com
- Rashmi Pandey is currently persuing masters degree program in Public Health from G.L.Gupta Institute of Public Health, University of Lucknow U.P. India PH-91-9889742990 E-mail:rashmip867@gmail.com