International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Security Challenges Associated with Big Data in Health Care System

N. Nagalakshmi, G. L. Anand Babu, K. S. Reddy, T. Ashalatha

Abstract: ***The way organisations handle, evaluate and leverage information in any sector has essentially altered with big data. Healthcare is one of the most promising fields where big data can be used to create a change. In this paper, as applied to the healthcare industry, we surveyed the state-of - the-art security challenges in big data, assessed how security issues arise in the case of large healthcare data, and discussed ways to address them. We concentrated primarily on the lately suggested anonymization and encryption techniques, their strengths and constraints, and envisaged future directions for studies.***

Keywords: Anonymization, Big Data healthcare, Encryption, Security.

I. INTRODUCTION

Big data in health care can imply improving patient results (improving services), better understanding of clinical services in demand and in which areas (creating efficiencies), and better payer efficiency (greater income). This implies that the healthcare company is pushing for more information and more data analysis. It also implies that there are important concerns about the proliferation of data in law and enforcement, especially any protected health information ("PHI") [1]. Any electronic PHI aggregated but not de-identified and evaluated in a fresh scheme or place under the Health Insurance Portability and Accountability Act ("HIPAA") Security Rule is subject to the same administrative, technical and physical safety checks as the scheme from which that information comes. Infringements of healthcare data have been cited as a top danger to the industry, according to the annual HIMSS Cybersecurity Survey. Ransomware that steals credentials for health care and malware are also significant fields of interest for information security professionals. Seventy-five percent of the 239 healthcare participants said their organisation had a major safety incident over the previous 12 months. The threat actor was identified by almost all of those organizations (96 percent). Thirty-seven percent of participants in the previous 12 months who encountered a safety incident said it was due to an online scam, such as phishing or spear phishing. Twenty percent of those surveyed ascribed the attack to a negligent insider, while another 20 percent said the problem was created by a hacker.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

N. Nagalakshmi*, Associate Professor, Department of IT, Anurag Group of Institutions, Hyderabad, Telangana, India.

G.L. Anand Babu, Associate Professor, Department of IT, Anurag Group of Institutions, Hyderabad, Telangana, India.

Dr. K.S. Reddy, Professor & Head, Department of IT, Anurag Group of Institutions, Hyderabad, Telangana, India.

T.Ashalatha, Associate Professor, Department of IT, Anurag Group of Institutions, Hyderabad, Telangana, India.

II. BIG DATA SECURITY IN HEALTHCARE

Hospitals cope with protected health information (PHI) and personal identifiable information (PII), both kinds of data must be protected at all costs by their safety schemes. Think of circumstances when a violation of hospital data reaches the headlines — it produces a enormous hype among nurses who are naturally worried about what occurred to their medical records, patient history, and information on payments. People responsible for protecting delicate medical information therefore have an significant job of maintaining due care in order to satisfy all regulatory data demands [2]. It is almost a "no-choice" situation as hospital safety is not something that can be taken lightly with excellent liberty of how to implement control measures for access. Security suppliers have implemented several alternatives at the front end. These include single sign-on, multifactor authentication, and patient identification to build confidence between customers, technology, and information across the healthcare environment. When they go to the doctor, most medical patients don't worry about their safety; they just expect their information and health to be taken care of using reliable safety measures. This is not always the case, unfortunately. As a consequence, it is essential for organisations to adopt data security solutions for healthcare that will safeguard significant assets while fulfilling compliance requirements for healthcare [3]. Technologies in use Different technologies are used to protect healthcare data's safety and privacy. The techniques most commonly used are:

A. Authentication

(a) Multifactor Authentication

Passwords can be an obvious objective for programmers, especially as new strategies, like password spray attacks and phishing arise that requires social engineering to exploit security system loopholes [4]. This is the place multifaceted authentication (MFA) techniques can step in to provide an additional level of safety of identification. In order to authenticate their identity and obtain access to a computer or device, MFA needs users to submit a mixture of variables — at least two. The variables fall into three classifications: that you are (like a biometric fingerprint), that you have (a portable phone) and that you know (a username and password). A typical mixture of two-factor authentication would be the user's username and password, as well as the user's smartphone's token code, Wright describes.

Many solutions also use biometric instruments that sense distinctive physical features such as fingerprint scanners or retina scanners. Moreover, according to a study published by HIMSS previously this year, phishing attacks are still one of the top threat performers for healthcare organisations, making the push away from passwords more urgent. "MFA greatly enhances safety, which is why you see a tremendous push to ensure that all high privilege accounts are not available without using some sort of MFA. In addition, MFA can even be a window into the password-free authentication globe. By deploying many of its own MFA alternatives internally, Microsoft has already been able to attain a pseudo-password less state for its customers.

(b) End-to-End Authentication

There are two issues facing healthcare cybersecurity – none of which can be solved with better technology alone. The first issue is that a critical aspect of our economy is healthcare. If you want to wreak havoc, the best place to begin might be to disrupt a country's healthcare system. In addition, extremely profitable are Electronic Health Records (EHRs). The second issue is that safety is a shared responsibility even within a single hospital. When a resource is commonly shared, the incentive of each party is to get as much benefit as possible while incurring the lowest possible cost at the same time. This issue is known as the "commons tragedy." Public Key Infrastructure (PKI) is the basis for end-to-end authentication. PKI authenticates users, systems and devices using digital certificates without the need for tokens, password policies or other complicated user-initiated variables. This decentralizes authentication and enables it to occur across different systems. In the healthcare industry, end-to-end authentication will not come to fruition until device manufacturers, hospitals, insurance companies, software suppliers and safety suppliers acknowledge their shared responsibility and start working together [5]. Cybersecurity is becoming much more of a pain point for those in the sector due to the increasing amount of exploits in healthcare. This pain causes some to behave and put in place greater safety.

B. Encryption

To safeguard their IT infrastructure, healthcare-based IT organisations need to leverage encryption techniques in two distinct ways. First, the information that crosses the networks must be protected. Next, the network infrastructure containing and distributing sensitive information must be protected. Encryption is a technique by scrambling information into an unreadable form to protect information. It is a systematic process of encoding that can only be reversed with the correct key. Encryption for the healthcare industry is a double-edged sword [6].

(a) Secure Socket Layer / Transport Layer Security (SSL/TLS) Encryption

SSL / TLS techniques provide the same advantages as Hedy's invention for radio communications for network communications (although the real processes are distinct). Secure Socket Layer / Transport Layer Security (SSL / TLS) encryption of in-flight private data is crucial for maintaining information privacy and integrity. The encryption provides security because the data is obfuscated by using complex mathematical algorithms that make it extremely difficult to decrypt the data without the appropriate keys or a huge amount of computing power. If the data is changed while on the network, the original information is lost as the encryption algorithm can not remove the data from the damaged communication. While the data is lost, the information's privacy is preserved and information can be resented across the guaranteed communication route. Many of the network infrastructure's assaults and threats are encrypted using SSL / TLS protocols. This burden inbound and outbound safety solutions with the job of decrypting communications, checking network traffic, and subsequently re-encrypting information. This is a big resource burden that can decrease these safety solutions ' efficiency by more than 80%.

C. Data Masking

Masking can scramble individual columns of data in various ways to make the masked data look like the original (keeping its format and type of data) but it is no longer sensitive data. Masking is powerful in maintaining aggregate values across a database as a whole, allowing the preservation of sum and average values within a data set while altering all the individual data elements. Masking plus encryption provides a strong mix for medical data distribution and sharing. Masked information should be quasi-real and realistic. It should comply with the same rules of company as the actual information. Using masked data in test and development settings is very prevalent because the data looks like "true" data, but does not contain any delicate information. It's a method that scrambles information, either a whole database or a subset. Masking is not reversible as opposed to encryption ; masked information is helpful for restricted purposes unlike tokenization. Several kinds of information masking are available:

- **Static Data Masking (SDM)** – pre-use masks information. Databases of non-production were NOT masked in real time.
- **Dynamic Data Masking (DDM)** – masks real-time manufacturing information
- **Data Redaction** – masks unstructured content (PDF, Word, Excel)

Each of the three information protection techniques (encryption, tokenization and information masking) has distinct advantages and works to address various safety problems.

D. Access Control

Upon authentication, consumers may enter an information system, but their access will sti8\l be regulated by an access control policy that is typically based on the privilege and right of each patient or trusted third party approved physician. Then it is a strong and flexible mechanism for granting user permissions. It provides advanced permission controls to guarantee that customers can only conduct the actions they are permitted to conduct, such as data access, work submission, cluster management, etc.

A number of alternatives to the safety and access control issues have been suggested. Hospitals still have some leeway over how they manage access control despite the need for powerful safety mechanisms and compliance with the law. Health care has a range of access control alternatives available which, as a general rule, rely on hospital expertise, architecture, and of course-budget [7].

E. Tokenization

Method to substitute non-sensitive placeholder tokens for sensitive information. These tokens are exchanged in relation databases and files with information stored. Data tokenization is the method of replacing sensitive information (such as Primary Account Numbers (PANs), Electronic Personal Health Information (EPHI) and Non-Personal Information (NPPI)) with a non-sensitive distinctive value. This non-sensitive value functions as a unique identifier for a sensitive record and is the "token." Tokenization can be used by many end customers as a constant value. This enables users to interact directly with the tokenized data without having to decrypt and re-encrypt data whenever they access the information. A reduction in the scope of the organization's compliance audits and validations is one of the most common benefits of using data tokenization. This also contributes to a reduction in the cost of regulatory compliance services and enables compliance attempts by the organization to concentrate more on other delicate and restricted data fields. Another advantage is the decrease of unauthorized access danger. There is no extrinsic meaning or value in the actual token, so there is very little risk associated with it [8]. Implementation of information tokenization also involves training staff on how to represent and use the information. Training, along with the daily use of tokens, can assist staff gain a deeper understanding of data protection's significance and thus better equip them to recognize control deficiencies that occur in other company fields.

III. RESULTS

The study discovered that healthcare organisations are also relatively uni-form in their safety risk assessment method. The following regions are included in their security risk assessments as stated by the participants:

- Policy and processes on cyber safety, paperwork (81.3% of participants)
- Network (74.7% of participants)
- Security knowledge and training (73.5% of participants)
- Physical safety (71.1% of participants)
- Inventory property (69.3% of participants)

Organizations can discourage assaults by establishing powerful filters to avoid third parties or unknown sources of information before they occur. All healthcare organisations need to get on the same page before healthcare cybersecurity can enhance. One way in which this can be achieved is by adopting a universal safety structure. We're not there yet, unfortunately.

Table 1: Summary of safety problems in Big Data

S.No	Method	Description
1	Multifactor Authentication	• Reduces password dependence • Improves safety • Increased accessibility of privilege accounts
2	End-to-End Authentication	• Proving the validity of all digital links • Public Key Infrastructure • Decentralizes authentication and enables for different applications
3	Encryption	• Data protection method by scrapping it in an unreadable form • Double-edged sword for the healthcare sector
4	Tokenization	• Replace sensitive information with a non-sensitive value • Reduce the cost of regulatory compliance facilities • Reduce the likelihood of unauthorized access
5	Access Control	• On the basis of the privilege and rights of each patient or trusted third party approved physician • Provide authorisation checks to conduct only the operations for which they are permitted
6	Data Masking	• Masking is efficient for keeping aggregate values across a database as a whole • Masked information should be realistic and quasi-real • Dynamic information masking (DDM) masks real-time manufacturing information

IV. CONCLUSION

Healthcare organisations are also relatively uniform in the process of assessing their security risk. The study discovered the top obstacles for cybersecurity remediation and mitigation were having the correct cybersecurity employees on employees and absence of economic resources. Organizations can discourage assaults by establishing powerful filters to avoid third parties or unknown sources of information before they occur. Big data security is considered the enormous barrier in this sector for scientists. We addressed some examples of effective associated job throughout the globe in this article.

REFERENCES

1. Rui Zhangand Ling Liu: ''Security Models and Requirements for Healthcare Application Clouds'' in IEEE 3rd International Conference on Cloud Computing, 2010

2. H. K. Patil and R. Seshadri, “Big data security and privacy issues in healthcare,” in 2014 IEEE International Congress on Big Data, 2014, pp. 762–765.
3. The Big Data revolution in healthcare, accelerating value and innovation – Peter Groves, Basel Kayyali, David Knott , Steve Van Kuiken –2013
4. D. Mondek, R. B. BlaÅ¿ek, and T. ZahradnickÃ¡, “Security analytics in the big data era,” in 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2017, pp. 605–606.
5. N. Khan, I. Yaqoob, I. A. T. Hashem, Z. Inayat, W. K. Mahmoud Ali, M. Alam, M. Shiraz, and A. Gani, “Big data: survey, technologies, opportunities, and challenges,” The Scientific World Journal, vol. 2014, 2014.
6. A. A. Cardenas, P. K. Manadhata, and S. P. Rajan, “Big data analytics for security,” IEEE Security Privacy, vol. 11, no. 6, pp. 74–76, 2013.
7. Challenges and Security Issues in Big Data Analysis. Reena Singh. Kunver Arif Ali. IJIRSET. Volume: 5. Issue: 1. January 2016.
8. M. Hagner. Security infrastructure and national patent summary. In Tromso Telemedicine and eHealth Conference, 2007.