Section 1

SECURITY AND PRIVACY CONCERN IN IoHT

1

Data Security and Privacy Concern in the Healthcare System

Ahuja Sourav

The NorthCap University, Gurugram, India

Abstract

Change is the new convention and standard for the world's health sector. Change in the collection, storage, and transfer of healthcare data through digitization is experiencing a fundamental as well as dramatic move in the executing, business and clinical archetype in the world's economy for the predictable future. This change is propelled by lifestyle changes and mellowing populations, the development of software applications, smart devices and gadgets, progressive treatments, and increased emphasis on care value and quality. All these will offer remarkable opportunities for improving and helping clinical objective decision, hence improving healthcare facilities and delivery, monitoring adverse events, and leading to a rise in the standards of treatment for diseases. In this chapter, we will read about developments in medical systems, requirements in medical data systems, and security and privacy concerns.

***Keywords*:** Healthcare system, security, privacy, eHealth data

1.1 Introduction

Medical data is being considered to improve health conditions, gain precious perceptions, and lower medical charges. However, increased privacy and security concerns are so irresistible that we are not able to take complete benefits of it with its available resources and facilities. The diverse technologies which generate data in medical fields are ever-increasing. The information generated is available at the central location and can be

Email: ahujasourav07@gmail.com

Kavita Sharma, Yogita Gigras, Vishnu Sharma, D. Jude Hemanth and Ramesh Chandra Poonia (eds.) Internet of Healthcare Things: Machine Learning for Security and Privacy, (3–26)

utilized for any organization including healthcare medical organizations, pharmaceutical manufacturers, hospitals, health insurance companies and all its sub-sectors [1, 15]. The flood of such large data has never been seen before in the healthcare sector. The success of all healthcare organizations needs to manage the analytical power of big healthcare data. This chapter covers privacy and security concerns, level of threat to information in healthcare organizations, security and privacy requirements in the medical field, existing solutions for security of healthcare data, future challenges and work in security and privacy of healthcare data, and privacy-preserving methods [5, 20].

Medical care is evolving as an existing infrastructure and is getting integrated with new technologies. New technologies including digital medical records and sensor-based monitoring of in-home patient remotely are at the current front row. Conventional paper-based medical reports are being digitized allowing patients to easily obtain their records and empowering healthcare workers to regularly monitor the health of the patients and save and protect their lives at the earliest. Specialized sensors and devices can be used at homes for remote patient monitoring [7, 9].

Patients remote monitoring offers powerful and flexible health monitoring through devices anywhere and at any time. The increasing demand, practicality, and convenience of digital medical care have brought various notable challenges for medical organizations, policymakers, patients, and medical centers [28]. A major challenge is dealing with privacy concerns which include patients' personal medical information which is kept online on third party clouds. It is the most significant and testing issue that medical care organizations must face in the future and preserve the patient's medical data and reduce the concerns of their privacy [6, 10].

Generally, the healthcare IoT devices layout is made up of the below-mentioned layers:

- Perception layer
- Network layer
- Application layer.

The perception layer collects healthcare data with the help of various devices. The network layer is composed of both wired as well as wireless systems and middleware. The network layer's major task is to transfer and transform the input data collected by the perception layer. The transport layer is well-designed. It improves transmission efficiency and reduces energy consumption. The transport layer also ensures privacy and security. The application layer provides healthcare services individually to users and

satisfies their needs by integrating medical information resources in accordance with the real situation of the service demand and target users [3, 19].

Data analytics in medical services has many benefits, and it displays extremely high potential for improving the medical care systems, yet it has many hurdles and challenges. Therefore, medical care data privacy and security concerns increase day by day. Also, a reactive, technology-centric method to determine privacy and security needs is not efficient by healthcare organizations to protect the medical centers.

The combination of these technologies will reduce costs and medical errors while enhancing the medical care standards by making them more sophisticated. To make new technologies and advancements in the field of medical sciences socially acceptable, related concerns of security need to be looked upon [24].

Security and privacy concerns of healthcare information are two required and essential concepts. Security of information means that data is collected, stored, and transmitted through secure channels. Data security guarantees validity, authenticity, and integrity. By privacy of data, we mean that the data can be retrieved by the users who have the authorization to access it [3, 18]. Based on different purposes and requirements, more reasonable protection strategies could be developed. The advancements and widespread use of healthcare sensors and devices provide a better guarantee to protect people's health. The advancements in the healthcare and medical field also put data privacy and security preservation under pressure. Security, as well as privacy, is the core consideration in the successful development of medical healthcare devices. New data storage and information systems along with new methods are required to avert breaches and other security happenings of delicate data to make efficient utilization of the large medical care information [11].

Both securities, as well as privacy, have impacted the way companies store, manage, and analyze sensitive data all over the world. Medical sciences along with information collected have enough capability to improve health results, predict the spread of diseases, get knowledgeable viewpoints related to the health of the society, prevent diseases in time, cut down the cost while also improving the standards of medical systems and raise the standards of living. Also, while securing user's privacy right, determining allowable uses of information is a difficult job [26].

To guarantee trustworthy and appropriate healthcare data, it is important to know the shortcomings of available systems and solutions in place. Privacy is typically known as the ability to preserve critical data about individuals related to medical reports. It aims to utilize and govern the private data of the users which includes forming new standards and

ensuring consent to ensure that users' data is being stored, transmitted, and used incorrect and secure course of action. Security is often known as preventing an unauthorized entry into the user's data. It also includes the availability and integrity of the data. It aims to protect sensitive information from deleterious attacks and information breach. Security is essential for protecting sensitive information, but it does not address privacy [1, 16].

1.2 Privacy and Security Concerns on E-Health Data

Security and privacy issues emerge from various vulnerabilities of information from the entire medical sector. It can happen between users, providers, and payers. The concerns of security and privacy in the healthcare sector are divided into two categories:

- Security and privacy worries about the improper release of data from individual organizations: Improper release of data from individual organizations can happen from following:
  - authorized users who access or spread data by infringement of organizational rules intentionally or unintentionally
  - The intruder who breaks into an organization's system.
- Privacy and security concerns regarding the systemic flows of data in the medical sector field and related sectors: It includes the open revelation of the patient personal medical data to parties that might be in opposition to the patient, and he might have thought of as breaching and violating the privacy of the patient [2].

1.3 Levels of Threat to Information in Healthcare Organizations

There are different types and levels of distinct types of organizational threats. The following threats are categorized as levels numbered from 1 to 5:

- Threat 1: *Insiders who do silly mistakes unknowingly which causes accidental revelations:* It is likely the most common source of privacy breaching which includes accidental

disclosure of personal information. It can happen in many following ways:

- Conversations between healthcare providers might be overheard anywhere
- Test results being noticed by the laboratory in-charge for familiarity in laboratory tests getting processed
- Computer screen left turn on which can lead to the passer seeing the information
- E-mail or fax sent to the wrong address
- Data left misclassified or misfiled.

- Threat 2: *Trustworthy insiders who misuse the privileges and access of data given to them:* These are the people who are given authorized access to medical records and infringe the trust corresponding to the privileges given to them. Medical workers, neither have the need nor the right to know the data but still, they are subject to curiosity in accessing information. There is no overall official statistics to indicate how many healthcare workers have accessed the medical records of others. Some reports indicate that medical care providers have seen medical records of others to estimate the chances of the sexually transmitted disease of others, details about their personal lives, with whom they were in close relation, with whom their former spouses were in relation. Medical staff tend to see the information about celebrities, politicians, actors, sports personalities, and other famous people which is potentially embarrassing health information, and the news regularly spreads out in the news.
- Threat 3: *Trustworthy people of an organization who intentionally investigate the data for some selfish motive:* The threat arises when that insider is authenticated to the partial data but not to the information he desires and through digital forensics and cyber tools gain unapproved access to that information.
- Threat 4: *Physical intruder who access the unauthorized data:* The invader has access to information but does not have authority and permissions for data access and to see the medical records. Many attackers pretend to be an authorized person by putting on a dress code and start using the system and get the desired health information of others.

- Threat 5: *Outsiders and employees who are vengeful who plans attacks to access the information they are unauthorized to, damage hospital assets, and obstruct medical operations:* The attacker does not have authority and physical access but still gain access to the data by purely technical means. Therefore, it is also called a pure technical threat. For example, the attacker breaking into a system from an external network applies cyber attacking tools and retrieves patient sensitive health records. Threat 5 is extremely dangerous. It is prominent that many healthcare organizations are heading towards using computer networks and distributed computing technologies in storing medical records, and so increasing the risk as patient records can be accessed through an external network. This is a dormant problem on the planet. It also has the risk of DOS attacks conducted by outsiders. DOS makes the attacked device useless for normal operations. An example of this can be, an outside attacker may access a sensitive medical information system and place a computer virus or Trojan horse into the computer that crashes the computer or erases important data healthcare records. Also, an intruder can burst your system with thousands of emails and could launch an e-mail attack. He/She can send a large number of emails in a span of less time, which is beyond the capacity of the mailing servers to process all emails. It leaves the computer and leaving the system futile for e-mail purpose [2, 29].

The invasion of privacy of the patient is a burgeoning concern in the field of information analytics. An incident reported quoted that Target Corporation sent baby care coupons to a teenage girl unbeknown to her parents. The incident raises an alarm over patient privacy and impels information analytics to consider privacy. Furthermore, it is essential for driving healthcare analytics to use privacy-protecting cryptographic encryption methods that permit managing algorithms of prediction on ciphertext while also preserving user specifications. However, the resource-constrained environment has the challenge to do the operation of exhausting resources while preserving privacy [1, 17].

For long, medical devices have been facing a lot of difficulties like, writing diagnoses on paper, making it difficult for doctors to access their patients' information. There is also a lack of space, time, and personnel to monitor patients. Development in the systems provide opportunities to better the

healthcare systems and minimize the problems. It would also provide personalized service. With the advancement, the privacy and confidentiality of patient's health records are at stake. Internet and network expose their personal information to hostile attacks. Attackers take advantage of the flaws and inadequacies of the system to collect the records of medical data and uncover it in front of the world. The traditional system does not face this problem as patients have to go into a medical laboratory or hospitals to receive their health reports which limit the count of people viewing the medical reports. Medical reports or records, once available online, open illegitimate doors for spammers and hostile attackers to capture and attack the records. Therefore, it is a big challenge to preserve information security as well as authenticity as opposed to the conventional medical care systems. While implementing EPRs as well as networking sensors, storage, authenticity, and access of information are the main challenges.

In the aspect of dealing with the challenges of electronic information, one should be aware of the following questions:

- Who is the owner of the information?
- What and what amount of information to be stored?
- At what place should the information be stored?
- Who all can be viewers of patients' medical records?
- Who can access this information without the patient's consent?

Rapid growth in the medical sector is enabling e-health records like EPRs in making the health reports available via the Internet to users/patients. and empowering the healthcare system being used for proper monitoring and analysis. In addition, the idea of monitoring patients remotely is turning into a practice by advancements in sensor networks [9, 14].

1.4 Security and Privacy Requirement

Although concepts of privacy and security are vital in medical healthcare systems, the maximum organization in medical sector does not spend many resources to preserve these constraints. Healthcare devices produce a highly sensitively large amount of real-time diverse data. In such a scenario, comprising data privacy and security and knockdown the security of the healthcare system or server could cause dangerous outcomes. At each step of information collection, storage, transmission, and analysis, the user's medical data exists. This can cause privacy breaches from multiple points of failure. While developing medical healthcare devices, equipment,

and network privacy and security systems, the below-mentioned needs should be taken care of.

- Data Integrity: Data integrity means that information satisfies linguistic standards without tampering through unauthorized and unacceptable methods. It comprises two levels of reliability as well as accuracy. It can be sub-divided into the following four categories:

  - Entity integrity
  - Domain integrity
  - Referential integrity
  - User defined integrity.

  These categories are managed by constraints, foreign keys, and triggers.
- Data Usability: Usability of data ensures the usage of data and data systems by authorized users. A large amount of data brings advantages and censorious challenges like deviant information and dirty data. Moreover, the usability of data is also ruined by data usage and exploitation due to unauthorized access.
- Data Auditing: Healthcare information audit access is a productive method to analyze resources utilization. Any abnormal events in the human body can be tracked by this measure. Also, cagey roles played by cloud service providers require rational auditing ways. Audit information usually includes those service providers, users, operation records, access, etc.
- Patient Information Privacy: Patient Information Privacy is the most crucial requirements in medical data privacy and security. Users' data is divided into below mentioned categories:

  - General patients' records
  - Critical information.

Critical information consists of information like identity information, sexual orientation, mental condition, sexual functioning, fertility status, infectious diseases, genetic information, and drug addiction. With such information about patient medical privacy, healthcare organizations need to

ensure that such information is not breached by unauthorized users. These organizations should also use cryptographic encryption to make sure if in case data is breached and intercepted, unauthorized users cannot understand the information and the patient's sensitive data is preserved [3, 30].

1.5 Security of Healthcare Data

Medical organizations collect, save, and manage a large amount of data for smooth functioning of the digital healthcare systems and providing efficient and proper healthcare to patients, but it has its downsides i.e., the shortage of technical support and minimum security. In recent times, healthcare systems are most vulnerable to data loss and exploitation which are extremely complicating. Attackers make use of data mining methods to get access to critical data and takes advantage of the vulnerability and causes critical information breach. The stakes for data privacy and security are continually raised but the ways for establishing and implementing security measures and defeating security controls remain a complex process and become more sophisticated gradually [27].

Therefore, keeping all the above-mentioned points in mind, organizations must establish and implement information security systems and solutions that will help preserve important healthcare assets which satisfy mandates of medical compliance.

1.5.1 Existing Solutions

Since healthcare devices and equipment are not so reliable concerning enough memory storage, processing power, and communication potential, they require an expanding, high-speed, and powerful big infrastructure for storage of information, data analysis and computing [3]. These concerns involving information access, storage, transmission, and processing are not new to the healthcare field [9]. At present, most healthcare organizations utilize cloud services from a third party and store the obtained healthcare records and deploy their applications on the servers. These devices are synchronized with their medical care tasks to the servers accordingly and store data on the cloud. Cloud services facilitate a propitious solution for managing ever-increasing and pervasive medical data by their flexibility and facility to utilize common resources and infrastructure in an omnipresent way [3].

The most widely used technologies to secure the privacy and security of medical data are as follows:

1) Authentication: Authentication confirms or establishes that the claims made about authentication and truthfulness are true. It serves essential functions in any organization like protection of the user identification, ensuring that the user claims of authentication are real and is the same person who is pretending to be an authentic user, and securing access to corporate networks. Data Authentication methods are used to confirm that the information sent is coming from the user it is being claimed to be from. Data authentication can have attacks like man-in-the-middle attacks. Many cryptographic protocols take measures to prevent man-in-the-middle attacks by including some form of endpoint authentication. A man-in-the-middle attack may be affected by persuading a user to agree with a certificate of digitalX 509. Therefore, while designing and building an end-to-end authentication solution, such scenarios must be considered. For instance, for providing communications network security such as over Internet protocols like transport layer security, secure sockets layer are cryptographic protocols that transport layer security and secure sockets layer are in use which encrypt end-to-end the network links segments at the transport layer. Browsing web pages, e-mails, faxing, real-time messaging, and voice-over-internet protocol (VoIP) is also controlled and managed by several versions of the protocols. To achieve authentication, hashing SHA-256 and the Kerberos mechanism can also be used. For authentication of servers, one can use SSL or TLS using a mutually trusted certification authority. To monitor all sensitive information, the algorithm of Bull Eye is useful. The algorithm of Bull Eye manages relationships between actual data and duplicated data and ensure information security. Both medical data obtained from the providers as well as the consumers' identities must be verified before giving access at each entry-level in healthcare systems. Various authentication protocols like passwords, digital signatures have been designed for better energy-efficient sensor networks, like the TinySec hash function [1, 5].

2) Encryption: Data encryption refers to preventing unauthorized access to critical data efficiently. It maintains as well as protects owner possession of information from the data center to the endpoint and into the cloud. It also includes electronic gadgets used by clinicians, administrators, and physicians. Encryption can be used to prevent theft of storage electronic devices and packet sniffing [1]. Cryptographic encryption is a technology for secure communication and data exchange in which the actual message is encrypted by the encryption algorithm into ciphertext by the agreed rules. The message is transferred from the sender of the message to the receiver through the public channel and the message is finally decrypted into plain text for the receiver to read. Encryption can be used to help

prevent eavesdropping and skimming and to make sure information security. Encryption can be done in both software and hardware. Using both software and hardware encryption guarantees the highest-level security.

Medical care organizations must ensure that the encryption algorithms are efficient. Healthcare providers should make sure that the cryptographic encryption method is efficient and easily usable by patients as well as medical professionals. It must also be simply and effectively able to include a large amount of new electronic medical records. Also, it is essential to take care that the number of keys holds by each party involved in the healthcare system should be minimized for easy maintenance and ease of usage. Encryption algorithms like RSA, AES, DES, triple DES, RC4, Rijndael, Blowfish, IDEA, and many more can be used for efficient encryption. It is a tough job to choose properly the appropriate encryption schemes to ensure storage security [1, 14]. Protocols of management play a very key role in the process of securing e-healthcare devices and their communication. But it is also evident that complex transmission protocols or encryption schemes can diversely impact the rate of transmission. It can even sometimes fail to do information transmission. Also, they are required to use valuable healthcare assets that are not readily available. Maintaining a balance between protection of security and the energy utilization of the system is tough and we need to be cautious while solving this as well as remain scientific with our approach.

3) Data masking: Masking replaces critical information with anonymous data. Actual data cannot be obtained from the masked anonymous data as this is not truly an encryption algorithm. It utilizes a de-identifying strategy of information sets which means masking personal identifiers like name, security number, and generalizing or suppressing quasi-identifiers like zip codes and date of birth. Therefore, masking data is a famous approach to live unidentifiable information. Data masking reduces the cost of securing a data deployment which adds up to its significant benefit. As secure data, masking decreases the need for using extra measures for security on that data after it is sent from a secure source into the platform and while it is in the platform [1].

4) Access control: Access control method is used to define the identity of a user by the data system. The predetermined strategy and policy prevent unauthorized users to access resources. Once the user is authenticated, he should be allowed to access an information system. But even after authentication, access control policy should still monitor their work depending on the access rights and privileges of every authorized user, if applicable. Role based access control is the supreme structure for advanced authorization control. Complexity and cost reduction of administration in huge

networked applications are the best possible results for the access control mechanism. This mechanism is immensely powerful and pliable to grant accesses to the medical sector. It offers advanced controls and authorization permissions to make sure that users can perform operations like information access, cluster administration, and job submission.

There are a lot of solutions proposed to address the issues of access control and concerns of security. Attribute based access control and role-based access control are the most famous access control models for solutions that are promising to control access control and security operations. In the medical system, attribute-based access control and role based access control have few drawbacks when used alone. We suggest adopting technologies in conjunction to satisfy the needs of access control with other fine-grained security techniques like encryption, and access control methods [1]. The encryption algorithms which can be utilized in access control strategies include:

- Symmetric key encryption
- Asymmetric key encryption
- Attribute-based encryption.

Cryptographic functions rely on keys. The cryptosystem's security is affected by the secret keys' length and its generation strategy. Hence, the life cycle of the security system is determined by the key management mechanism for the cryptosystem. Attribute-based encryption has slowly become a mainstream mechanism due to the ascendable key management and pliable access control policies. Medical data exchange provides electronic medical data with explicit data exchange authorization in an ascertainable manner. Currently available non-cryptographic methods for authentication exhibit several drawbacks. It lacks a reliable and secure strategy for enforcement of access policy in medical information systems. Cryptographic encryption approaches, on the other hand, are too costly, complex, and restricted in identifying policies [3].

5) Monitoring and auditing: Auditing and monitoring security are collecting information to investigate events to find the invaders and intrusions. Monitoring and keeping a record of activities that user does and keeping a log of every access in the medical system in chronological order is called an audit. Cloud servers are not reliable and cannot be trusted completely. The consistency, integrity and availability of healthcare information saved on the cloud might be compromised if information corruption or deletion occurs without the consent of an authorized person. The information guidelines are typically specified by the authorized person for security purposes. This ensures no direct contact of the service provider with the

information. The trusted third party comes up with a high reputation for providing unbiased audit reports. In addition, these trusted third parties can be established to preserve the licit benefits of users on the cloud and enable the accountability of cloud service providers properly. These third-party auditing also has few issues which comprise batch auditing, dynamic auditing, and performance metric auditing. Various auditing methods have been presented in the past few decades which includes logistic regression and support vector machine. These auditing methods can detect suspicious access. However, large-scale promotions are restricted by depending heavily on judgments which are expected and predefined tags. In the present time, unsupervised approaches attract more attention to the healthcare sector [1, 3].

1.5.2 Future Challenges in Security and Privacy in the Healthcare Sector

Developers while developing healthcare devices or programs, to get a better balance in the output, considers various factors and their impact which include security and privacy. For achieving a better security environment, various obstacles need to be addressed and require special attention.

- Insecure Network: Many devices and software rely too much on wireless networks like Wireless Fidelity, as they are convenient and have a lower cost. These systems are supposed to be vulnerable to many attacks, intrusions and penetrations which includes MITM attacks, unauthorized router access, spoofing, brute-force attacks, traffic injections, and DOS attacks. Also, untrusted networking paths that have not been certified possess a greater challenge than most free wireless networks in a public place.
- Lightweight Protocols for Devices: specific policy and proxy rules should be followed by the lightweight and cheaper devices and software/hardware sensor-based applications to deliver services. In the current situation, we must apply comparatively expensive solutions if we wish to deliver higher grade security to the sensors. It is a conflict in the medical healthcare system. The main task of security protection is to develop different levels of security protocols like lightweight security protocols according to application scenarios in the future.
- Data Sharing: Information Island phenomena is a big concern, despite the rapid advancement of medical healthcare

data systems. It is difficult to manage the healthcare data gathered as the level of the information collected from devices of different companies vary diversely. However, the inevitable trend of the future constitutes data collaboration and sharing within heterogeneous systems of medical care. Medical care systems determine the privatization of patient data. Privacy and security enhancement can be done with a hierarchical security model by collaborating general information policies to amalgamate different information which might provide better comprehensible data and can improve the privacy and security of the system [3].

1.5.3 Future Work to be Done in Security and Privacy in the Healthcare Sector

Though there are techniques that are used to cure the privacy and security of healthcare information with the above-mentioned technologies, there is always room for improvement.

- Defining proper attributes for role-based accesses: To put the healthcare system in place, logical attributes need to be established and implemented properly for the role based access. It can either be static rules or dynamic rules based upon what is appropriate as per the requirements.
- Policy development: old policy should be alerted and updated policy should be formulated that can cope with jurisdiction across state. HIPAA rules of privacy are equipped with some groundwork, but better-refined policies should be created to create rules that are clear and also users can bank upon. The step towards EPRs (Electronic Patient Records) and the growing amount of healthcare information will be generated remotely because networking sensors create the ability and need to transfer huge amounts of information rapidly. This entails a complete set of rules and procedures that safeguard a user's security as well as privacy. Many patients are ill-informed or are unsure regarding their privacy rights related to healthcare information. In the current setting as more healthcare information becomes available online and can be transferred easily, the increase in healthcare data will also

broaden the uncertainty of users many folds unless comprehensive guidelines and policies are well determined and well informed [22, 25].

- Rules on patients' privacy at home: Guidelines and policies are required to be formulated that will be regulating what sensor information collected necessitate, where it is stored and who will have control over it. Policies are required to be formed which comprises that the user should have complete control of the information which is stored on a central monitoring storage station or patients should at most have fragmented control. Standards and guidelines should also be established over access controls available to patients for handling their health data and reports.
- Data mining rules and technological measures: Rules, regulations and technological measures of data mining includes which person has the authority to analyze what type of information and the rules on collected anonymous information. If some form of automatic technology is feasible, the appropriate technological methods for ensuring these controls are intact [9].

Security and privacy in medical sciences is a complex undertaking. To summarize the overall problem of security in medical sciences, the following points can be considered:

- An extensive surface area for an invasion
- The cloud services transformation
- The need for data privacy
- The shortage of talent compounds.

In the near future, we should plan to address these issues through innovation and advanced technologies in healthcare that can offer a consistent way to work with electronic healthcare data and to establish policies by supporting a better platform. We should even utilize machine learning algorithms to easily detect malware, Trojans, and other cyberattacks. These technologies can be ideal for medical science to grow and will provide healthcare organizations long-term durable firmness across the security spectrum and privacy needs. Invader detecting and preventing methods are tricky to have on the complete traffic of the network. To resolve this issue, a security surveillance infrastructure is built which will examine DNS network traffic, honeypot data, IP flow data,

and HTTP network traffic. The optimal solution is to use data correlation schemes which include storing and processing data in distributed sources. At this stage, to recognize whether the domain name, the data packet or network flow is spiteful, three likelihood metrics have been calculated. An alert pop up in the detection system depending on the score obtained by this calculation, or the process terminates by the prevention system [1].

1.6 Privacy-Preserving Methods in Data

Newer privacy preservation methods are helpful to secure the privacy of the data. They are listed below as follows:

1. De-identification: It is a conventional approach to forbid the revealing of sensitive private data by declining the data that could recognize the patient by the following:

    - the removal of patient's identifiers
    - statistical method where it is verified that patient's enough identifiers have been deleted.

An external may still obtain information help for de-identification. Therefore, de-identification does not guarantee the protection of data-privacy. We can develop more efficient privacy-preserving algorithms which shall be feasible to weaken the threat of re-identification. Few concepts have been introduced to enhance this traditional technique like k-anonymity, t-closeness, and l-diversity.

- k-anonymity: This approach works on the principle that the k value is inversely proportional to the re-identification probability i.e., the more the k, the lesser will be the re-identification probability which can result in deformations of data leading to great information leakage due to k-anonymization. Also, excessive anonymization should not be done as it reduces the usefulness of the disclosed data for the recipients as some of the investigations may produce partial and untrue results or sometimes even becomes impossible. In k-anonymization, the sensitive identifier like a disease will be revealed. The two methods to accomplish k-anonymity for a 'k' are Suppression and Generalization.

  - Suppression: This method includes an '*' which could replace some attributes' values. Some or all the values of a column can be supplanted by '*'.
  - Generalization: This method includes replacing the individual values of the attribute with a broader category.

- L-diversity: we moved towards the l-diversity strategy of data anonymization to reflect upon the actual usefulness of the data. It is one of the batch-based anonymizations. It is used to protect data sets privacy by minimizing data representation granularities. This model has the features of Recursive, Entropy, and Distinct. It is the extended version of the k-anonymity. K-anonymity minimizes the granularities of information display that any data maps onto at the minimum k different data records by utilizing different methods which includes suppression and generalization. The k-anonymity model had few weaknesses that safeguarded identities 'k' individuals are not equal to the respective critical data values that were suppressed or generalized. This weakness is handled by the l-diversity model. This method depends upon the range of sensitive attributes which is the drawback of this method. If want to make data L-diverse fictitious data to be inserted even though delicate accreditation has not as much as different attributes. This counterfeit data may result in problems amid analysis, but it will help improve security. This results in the l-diversity technique depending on unbalanced and similitude strike and hence cannot avoid divulgence of the element.
- T-closeness: This can be considered as further advancement and enhancement of l-diversity batch-based anonymization. Its imitation has the features of equal and hierarchical distance. It extends the l-diversity model by considering the arrangement of data values for the particular aspect and treating the values of an attribute distinctly. The major benefit of the t-closeness approach is that it seizes disclosure of attribute. The major drawback is that the re-identification probability rises as the diversity of data rises.

2. HybrEx: The hybrid execution imitation is a model in cloud computing that deals with concealment and solitude. It uses public clouds only in the following cases:

- Organization's insensitive data
- Computation grouped as public.

Computation grouped as the public is when the organizational body announces that solitude and concealment risk is nil in exporting the information and computing it by utilizing public clouds. But the model executes their private cloud for an organizational body's private, delicate data and calculation. Moreover, the solicitation itself also gets divided and performs in both public as well as private clouds when a supplication requires access to both public as well as private clouds. It provides integration with safety and considers data susceptibility before a job's implementation. The drawback of HybridEx is, it does not manage the key that is executed in the map phase at private and public clouds. Another disadvantage of it is, it works with the cloud only as an opponent [1, 10].

There are four categories in which HybrExMapReduce utilizes both public and private clouds and enables new kinds of applications. The following are the four categories of HybrExMapReduce:

- Map hybrid: The diminished phase is implemented in only one of the clouds while the map phase is implemented in both private as well as public clouds.
- Vertical partitioning: Maps and minimizes tasks that are implemented using public data as the input in the public cloud and save the result in the public cloud after shuffling intermediary data amongst them. A similar task is executed with private data in the private cloud that the works are prepared in segregation.
- Horizontal partitioning: In this, the reduce phase is performed in a private cloud when the map phase is performed only in public clouds.
- Hybrid: The map phase and the diminished phase are executed on both private and public clouds. It is also possible to do data transmission among the clouds.

3. Identity-based anonymization: Identity-based anonymizati tion is a kind of data sanitization. The purpose of identity-based anonymization is privacy protection. It is the process in which the persons described by the data remain anonymous either by cryptographic encryption or eliminating personally identifiable data. The main drawback of this technique is that while

protecting the identities, it involves privacy protection combining anonymization and large data approach to examine usage data [1].

Medical Healthcare equipment is the collection of Internet of things devices coupled to the information servers, to execute the services and processes that expand and expand health protection. Heath care IoT based equipment has come out as an upcoming approach for electronic medical care that stores essential vitals of the patient's body, stores them, keep a check on their pathological details with small gadgets that are easily wearable or implantable sensors, and analyzes the patient's condition to stay pre-active in critical conditions. Advancements in healthcare have proved great prospect in improving the medical care and lifestyle of human beings and support many gadgets from embedded healthcare instruments to the wireless body area network [3, 12]. The digitization of healthcare systems and medical records has set a new wave and paradigm shift in the healthcare industry in recent times. Thus, there is a sheer rise in the amount of data in the healthcare industry which has increased complications, variety, and timeliness. The terminology "big data" has been in a lot of use in digital healthcare units which indicates the cluster of complex and big sets of information, which is far more superior to existing computing, storing data, and communicating capabilities of already existing methods or systems. In the previous twenty years, the hiked cost is at a frightening rate. Estimated medical care expenditure is now at 17.6% of GDP. Here, big data emerges as a possible solution as medical care officials strive for every feasible way to reduce the costs while enhancing the healthcare system, its delivery, and functioning. However, there is a makeable increase in privacy and security matters with the acquiring of big data in medical care systems. At the outset, data centers are made with different levels of safety where patients' Information. Also, most of the information centers which store healthcare data in large volumes have HIPPA (Health Insurance Portability and Accountability Act) certification which aims at ensuring safety policies and procedures. HIPAA certification does not guarantee patient record safety [4, 21].

There is also an increase in the popularity of healthcare cloud solutions which has increased the complexities in safeguarding big, distributed SaaS solutions. There is an invasion of huge data sets with varying data sources and formats that adds an extra burden on data storage and processing. In the burgeoning healthcare industry, examining security threats and foreseeing risk sources in an actual environment is of maximum need. The flood of attacks ranging from stealthy malware to Distributed DoS is a

matter of concern in the healthcare industry. It is hard to foresee without taking human cognitive behavior into account which is affected by the rise in social engineering attacks and the risks associated with it [4, 8]. In today's IoT based environment, it is incredibly challenging to implement security and privacy in resource-constrained networks and the complexity will keep expanding with the rapid expansion in the count of Internet of Things-based devices. Typical symmetric and asymmetric distribution of cryptographic keys cannot be expanded to more than a billion Internet of Things-based devices. Therefore, the new solutions key management between disparate networks like IP networks and the Internet of Things were scalable and crucial for the integration of the Internet of Things and big data in a cloud-based environment. It will lead to the seamless interoperability of data of healthcare systems [13, 23].

As the healthcare industry leverages emerging big data, analytics of security analytics will be at the utmost priority of any healthcare industry related design leveraging on expanding technologies in the context of big data. Big data along with SaaS solution will make better knowledgeable decisions hosting Protected Health Information (PHI). Moreover, real-time data analytics can be a challenge in the administration of such systems [4]. If security and privacy concerns are intact, medical IT experts can detect risks on critical patients immediately and take pre-emptive steps to control the deteriorating health condition of an individual [18].

1.7 Conclusion

Big healthcare data can be useful for the advancement of medical science. But in driving such technologies, healthcare, security, and patient privacy are paramount. Managing organizations are not willing to share massive their vast medical records for centralized analytics. Securing patient data management is inevitable when medical cloud networks link and generate large amounts of data [4]. It is especially important for the successful digitization of health systems, but it can be utilized only if concerns of privacy, as well as security, are addressed. As it happens, medical care clouds with huge healthcare information will become notable. Additionally, immediate processing will impel proactive healthcare in a secure and privacy-preserving environment [4]. In recent times, with the increased popularity of healthcare innovations and analytics, new privacy laws need to be put in place to demonstrate the entire process involved in analyzing large medical data and to protect patient privacy. Before performing such analytics

and applying algorithms on health information, "informed consent" from patients is duly required [20].

At present, to increase the generation of the huge amount of information as well as to improve the quality of medical services, different medical healthcare devices and software are being developed. The importance of this huge amount of medical care information is self-evident. In future related research, the fact that how to effectively protect data security, privacy and prevent data breaches will occupy an important position at all stages of data flow [3, 18]. Attention has been given to medical healthcare systems. The policies, standards and technical guidelines are still under progress in the healthcare system intending to protect lives at the earliest.

References

1. Abouelmehdi, K., Beni-Hessane, A., Khaloufi, H., Big healthcare data: Preserving security and privacy. *J. Big Data*, 5, 1, 1, 2018.
2. National Research Council, Privacy and Security Concerns Regarding Electronic Health Information, in: *For the Record Protecting Electronic Health Information*, National Academies Press (US), 1997.
3. Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S., Wang, G., Security and privacy in the medical internet of things: A review. *Secur. Commun. Netw.*, 1–8, 2018.
4. Patil, H.K. and Seshadri, R., Big data security and privacy issues in healthcare. *2014 IEEE International Congress on Big Data*, IEEE, 2014.
5. Abouelmehdi, K. *et al.*, Big data security and privacy in healthcare: A Review. *Proc. Comput. Sci.*, 113, 73–80, 2017.
6. Ren, Y., Werner, R., Pazzi, N., Boukerche, A., Monitoring patients via a secure and mobile healthcare system. *IEEE Wireless Commun.*, *17*, 1, 59–65, 2010.
7. Meingast, M., Roosta, T., Sastry, S., Security and Privacy Issues with Healthcare Information Technology. *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, NY, pp. 5453–5458, 2006.
8. Rodrigues, J.J.P.C. *et al.*, Analysis of the security and privacy requirements of cloud-based electronic health records systems. *J. Med. Internet Res.*, 15, 8, e186, 2013.
9. Kumar, J.S. and Patel, D.R., A survey on internet of things: Security and privacy issues. *Int. J. Comput. App.*, 90, 11, 20, 24–25, 2014.
10. Chen, D. and Zhao, H., Data Security and Privacy Protection Issues in Cloud Computing. *2012 International Conference on Computer Science and Electronics Engineering*, Hangzhou, pp. 647–651, 2012.
11. Wilkowska, W. and Ziefle, M., Privacy and data security in E-health: Requirements from the user's perspective. *Health Inf. J.*, 18, 3, 191–201, 2012.

12. Al Ameen, M., Liu, J., Kwak, K., Security and privacy issues in wireless sensor networks for healthcare applications. *J. Med. Syst.*, 36, 1, 93–101, 2012.
13. Zhang, Y., Zheng, D., Deng, R.H., Security and Privacy in Smart Health: Efficient Policy-Hiding Attribute-Based Access Control. *IEEE Internet Things J.*, 5, 3, 2130–2145, 2018.
14. Rahim, F.A., Ismail, Z., Samy, G.N., Information privacy concerns in electronic healthcare records: A systematic literature review. *2013 International Conference on Research and Innovation in Information Systems (ICRIIS)*, Kuala Lumpur, pp. 504–509, 2013.
15. Burghard, C., *Big data and analytics key to accountable care success*, pp. 1–9, IDC Health Insights, Framingham, 2012.
16. Fernandes, L.M., O'Connor, M., Weaver, V., Big data, bigger outcomes. *J. AHIMA*, 83, 10, 38–43, 2012.
17. van der Linden, H. *et al.*, Inter-organizational future proof EHR systems: A review of the security and privacy related issues. *Int. J. Med. Inform.*, 78, 3, 141–160, 2009.
18. Gong, T., Huang, H., Li, P., Zhang, K., Jiang, H., A Medical Healthcare System for Privacy Protection Based on IoT, in: *Proceedings of the 7th International Symposium on Parallel Architectures, Algorithms, and Programming, PAAP '15*, pp. 217–222, 2015.
19. Hu, J.-X., Chen, C.-L., Fan, C.-L., Wang, K.-H., An intelligent and secure health monitoring scheme using IoT sensor based on cloud computing. *J. Sensors*, 2017, Article ID 3734764, 11 pages, 2017.
20. Groves, P., Kayyali, B., Knott, D., Kuiken, S.V., *The 'big Data' Revolution in Healthcare*, McKinsey & Company, New Jersey, New York, 2013.
21. Lee, W.-B. and Lee, C.-D., A Cryptographic Key Management Solution for HIPAA Privacy/Security Regulations. *IEEE Trans. Info. Tech. Biomed.*, 12, 34–41, 2008.
22. Meingast, M., Roosta, T., Sastry, S., Security and privacy issues with healthcare information technology. *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006.
23. Jiang, D. and ShiWei, C., A Study of Information Security for M2M of IoT. *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, pp. 576–579, 2010.
24. Meingast, M., Roosta, T., Sastry, S., Security and privacy issues with healthcare information technology. *Conf. Proc. IEEE Eng. Med. Biol. Soc. 2006*, vol. 1, pp. 5453–5458, 2006.
25. Choi, Y.B., Capitan, K.E., Krause, J.S., Streeper, M.M., Challenges associated with privacy in healthcare industry: Implementation of HIPAA and the security rules. *J. Med. Syst.*, 30, 57–64, 2006.
26. Corritore, C.L., Kracher, B., Wiedenback, S., Marble, R., Foundations for trust for e-Health, in: *Emerging Healthcare Systems*, M. Ziefle and C. Röcker (Eds.), pp. 49–75, IGI Global, Hershey, PA, 2013.

27. Kargl, F., Lawrence, E., Fischer, M., Lim, Y.Y., Security, privacy and legal issues in pervasive ehealth monitoring systems. *7th International Conference on Mobile Business icmb*, vol. pp. 296–304, 2008.
28. Milenkovic, A., Otto, C., Jovanov, E., Wireless sensor network for personal health monitoring: issues and an implementation. *Comput. Commun.*, 29, 2521–2533, 2006.
29. Samy, G.N., Ahmad, R., Ismail, Z., Threats to Health Information Security, in: *2009 Fifth International Conference on Information Assurance and Security*, pp. 540–543, 2009.
30. Parks, R., Chu, C.-H., Xu, H., Healthcare Information Privacy Research: Issues, Gaps and What Next?, in: *Americas Conference on Information Systems (AMCIS) 2011 Proceedings*, 2011.