A Privacy-preserving Framework for Smart Context-aware Healthcare Applications

Muhammad Ajmal Azad[a], Junaid Arshad[b], Shazia Mahmoud[b], Khaled Salah[c], Muhammad Imran[d]

[a]*Department of Computer Science and Mathematics, The University of Derby, Coventry, United Kingdom*
[b]*School of Computing and Engineering, University of West London, London, United Kingdom*
[c]*Department of Electrical and Computer Engineering, Khalifa University, Abu Dhabi, UAE*
[d]*College of Applied Computer Science, King Saud University, Saudi Arabia*

Abstract

Internet of things (IoT) is a disruptive paradigm with wide ranging applications including healthcare, manufacturing, transportation and retail. Within healthcare, smart connected wearable devices are widely used to achieve improved wellbeing, quality of life and security of citizens. Such connected devices generate significant amount of data containing sensitive information about patient requiring adequate protection and privacy assurance. Unauthorized access to an individual's private data constitutes a breach of privacy leading to catastrophic outcomes for an individuals personal and professional life. Furthermore, breach of privacy may also lead to financial loss to the governing body such as those proposed as part of the General Data Protection Regulation (GDPR) in Europe. Furthermore, while mobility afforded by smart devices enables ease of monitoring, portability and pervasive processing, it also introduces challenges with respect to scalability, reliability and context-awareness for its applications. This paper is focused on privacy preservation within smart context-aware healthcare with a special emphasis on privacy assurance challenges within the Electronic Transfer of Prescription (ETP). To this extent, we present a case for a comprehensive, coherent, and dynamic privacy-preserving system for smart healthcare to protect sensitive user data. Based on a thorough analysis of existing privacy preservation models we propose an enhancement for the widely used Salford model to achieve privacy preservation against masquerading and impersonation threats. The proposed model therefore improves privacy assurance for cutting edge IoT applications such as smart healthcare whilst addressing unique challenges with respect to context-aware mobility of such applications.

Keywords: Smart healthcare, e-healthcare, context-awareness, security and privacy, context-aware mobility

1. Introduction

Internet of things (IoT) is a disruptive paradigm with applications across diverse domains including healthcare, manufacturing, transportation and retail. The smart medical devices market is expected to reach revenue of $25 billion revenue by the year 2025 with such smart connected wearable devices envisaged to be widely used to achieve improved wellbeing, quality of life and security of citizens. In addition to their ability to support real-time continuous monitoring of patients' vital signs, such devices also enable context-aware mobility vital to improving overall quality of medical care provided such as those illustrated by Bhattacharyya et al.[1].

Smart interconnected devices generate large amount of data that is used for analysis and improved recommendations to the patients and overall healthcare system. However, this data represents patient behaviour and includes medical history and therefore requires adequate protection and privacy assurances. Within this context, healthcare data refers to any health-related information that is relevant for making the decision about the individual's health. The collected data can be used for prevention, treatment, cure, health promotion, self-care, and wider public health activities. This data provides an insight into an individuals wellbeing, can affect treatment procedures, and helps study emerging trends in diseases and to modify treatment protocols [2, 3, 4].

There are different sources for healthcare data such as administrative data, patient medical record (PMR), patient surveys and the standardized clinical data [5]. The electronic patient records include a number of patient related information, such as physicians notes, patient comments, and questions, clinical trials, patient surveys etc. Specialist doctors, physicians, nurses, patients, can all have access to this record over the Internet. With the digitization of all these data records, the volume, variety, and veracity of the healthcare data have increased, making it complex and diverse, therefore requiring emerging paradigms such as Big Data to facilitate the efficient processing to extract useful knowledge [6, 7].
However, the digitization and outsourcing of data to computing rich organizations make it accessible to wider user domains, on-demand with minimal effort. Such patterns of data accessibility introduce new challenges with respect to security and privacy due to the sensitive and confidential nature of the personal data [8, 9, 10, 11]. Furthermore,

while mobility afforded by smart devices enables ease of monitoring, portability and pervasive processing, it also introduces challenges with respect to scalability, reliability and context-awareness for its applications.

The European Union has introduced a General Data Protection Regulation (GDPR) [12] regulation for mandating organizations to implement data protection measures. The compliance of GDPR regulations has brought an added emphasis to the protection of security and privacy, especially for the personal data. Various measures have been adopted for protecting personal data of individuals from the misuse. These include data anonymization [13, 14, 15] i.e. anonamyzing the individual's personal information and sensitive fields, or using the cryptographic tools [16, 17, 18, 19, 20, 21, 22] to protect the confidentiality, integrity, and privacy of individual data. Furthermore, authentication mechanisms are also deployed to verify the identity of a user attempting to access the data. These include password and certificate-based systems, biometric-based access system and multi-factor authentication [23, 24]. A Salford model is used along with the paper prescriptions and barcodes so that the patient should still avail the health services in the event of network or service failure [25]. Since health care workers are involved in different roles and responsibilities, their access to the system can be based on their job description [26, 27, 28].

In this paper, we performed a comprehensive study on the security and privacy challenges related to the electronic healthcare data with the patient portals. In particular, we discuss real-life healthcare scenarios i.e. Electronic Prescription Transfer [29], access to a patient personal data in an emergency and the use of mobile technologies to access the healthcare data [30, 31]. Further, we analyze the potential security and privacy properties of the existing systems and propose an enhanced version of the Salford model to improve the security and privacy properties. The enhanced model does not increase the complexity of the system and does not incur additional computational and system resources. Additionally, we analyzed the security and privacy properties of the enhanced model with the state of the art solutions and provide recommendations on involving patient with specific attention to consent and approval due to the sensitivity of the data involved.

Rest of the paper is organized as follows: Section 2 provides a discussion of the primary sources of healthcare data. Section 3 identifies security concerns related to different healthcare data sources. Section 4 presents an in-depth analysis of the factors affecting security and privacy in healthcare applications followed by an analysis of the unique challenges within these applications in section 5. Current security and privacy measures are presented in section 6. Section 7 discusses potential scenarios within healthcare applications and proposed improvement for the Salford model in Section 8. Section 9 discusses future recommendations. Section 10 concludes the paper with a discussion on open challenges to improve the security and privacy of healthcare data.

2. Healthcare Data Sources

Healthcare sector is diverse and multidimensional involving entities such as patients, doctors, lab technicians, administrative staff, health centers, insurance companies, and pharmacies. These entities interact with each other generating unique but interrelated data. This section focuses on the major sources of healthcare data.

2.1. Administrative data

Any interaction between a person and healthcare system produces data categorized as *administrative data*. These interactions include but are not limited to visiting a General Practitioner (GP), diagnostics reports, medical prescriptions, admission/discharge from a healthcare facility, referrals to other GPs or specialists and prescribed medical tests for further analysis. Administrative data are collected through claims, enrollments, procedures, and diagnostics tests and is primarily used for administrative and billing purposes but can be used to study the effects of procedures, quality of the healthcare services, side effects of the drugs and costs of healthcare services.

The administrative data can also be used to analyze pharmacy services, drug utilization and effects, and drug safety and effectiveness. Research based on administrative data must be interpreted in the context of changes in drug availability over time. Introduction of a new drug to the market and changes to reimbursement criteria or health coverage can significantly affect drug utilization and patients classification.

Typical elements of the administrative data are types of services, units of services utilized, cost of diagnosis, codes of procedures, location code and amount billed with current balance. As these criteria and codes are consistent, they can be easily accessed and shared globally [32]. Furthermore, as it is primarily used for billing and collection purposes, it contains limited medical information and therefore, has limited significance for public reporting.

2.2. Patient health records

A patient health record is a list of medical care and services being used by a patient. Over the last two decades, most of the medical data produced are in a digitized form which facilitates ease of access, storage, and sharing. Furthermore, it also makes the data convenient and cheaper to use for healthcare service improvement, quality control and reporting of healthcare sector progress. Patient health records typically act as a communication tool and reference for treatment, long-term care, and management of chronic medical conditions [33, 34, 35]. The prime benefit of this source of healthcare data is its richness in clinical data and overall credibility. However, the limitations of medical record are its complexity and compilation of

data in different formats retrieved from different healthcare sources. Additionally, since a significant segment of the data is in paper format, it requires expertise, time and effort to extract required information.

2.3. Patient surveys

Patient surveys are periodic and are used to gather strategic information about patient experiences with the healthcare services. It not only gives an overall picture of the patient satisfaction with healthcare level of services but also gives the progress and direction in which healthcare system is heading highlighting areas for improvement. It includes treatment, medical care and use of healthcare facilities by the patients and their level of satisfaction with the quality and effectiveness of the healthcare services [36, 37]. These surveys are usually administered through the use of phone, post or Internet-based tools. Patients satisfaction and overall perception of the healthcare services are the basis of patient-oriented services since the ultimate indicator of a healthcare system is patient contentment. Patient surveys are a useful source of healthcare data as patients are the best source to evaluate the current condition of the healthcare sector. Results of the surveys are easier to understand and share as they are provided by the patients themselves. However, potential shortcomings of this source of data include the expense, time consumption, and significant impact of the formulation of questions and sample population size on the results [36].

2.4. Comments from individual patients

Any informal information gathered from or provided by patients as opposed to carefully designed efforts is termed as *comments from individual patients*. This source of healthcare data is becoming increasingly common and gaining influence as the Internet and social networking sites are making it easier for individuals to share their expressions, experiences, and preferences. Therefore, feedback from individual patients is becoming an increasingly popular ingredient regarding the performance of healthcare services, health plans and healthcare providers especially physicians [5]. It is a powerful tool as the knowledge available from individual patient's feedback is immensely rich as compared with statistical information.

Furthermore, this source provides patients a platform to communicate their expressions and feedback. However, major drawback of this source is its integrity and its susceptibility to being influenced by emotions. Furthermore, the results are not impartial as comments are not collected systematically. Another issue with this source is the fact that it does not represent the entire patient population as ratings of a facility, GP or service are random and can be inconsistent across the patient population.

2.5. Standard clinical data

Standard clinical data represents data reported by facilities such as nursing home, non-profit organizations involved in vaccination and certain health agencies regarding the health and treatment given to the patients on a regular basis. This data provides details of healthcare information and can be used for quality control, quality assurance and assessment of the health agencies. This allows for collaborative research, large-scale analytics and sharing of sophisticated tools and methodologies and therefore important because it ensures traceability of the data and optimized data flow. Also, these statistical analyses require limited modifications, minimal oversight, and resource allocation. It also involves simple mapping algorithms ensuring consistency between *Operational Data* and *Submitted Datasets* [38]. A major limitation of this source is that it may not address all topics of interest.

3. Security Challenges for Healthcare Data

Healthcare data is diverse and can be categorized into different types as detailed in section 2. Although different types of healthcare data are interrelated, they are unique and thus have unique security and privacy issues based on their source and type of information they possess [39, 40, 41]. In this section, we present a discussion regarding security and privacy challenges for different healthcare data sources.

3.1. GDPR Regulations and Patients Health Data

The European Union has forced it member countries to employ the General Data Protection Regulation (GDPR) on 25 May 2018. The GDPR is the European Unions new law for ensuring the protection of citizens private data and is replaces EU existing law i.e. Data Protection Directive, which has been in effect since 1995. The GDPR regulations imposes wide range of requirements and regulations on the organizations that used to collect and process personal data of citizens. These requirements includes transparency, fairness and lawfulness in handling and processing the personal data of citizens. Furthermore, the law also limits organizations to process the personal data with the consent of users and for the legitimate purposes. It also requires that organization must ensuring the security, integrity and confidentiality of personal data by imposing a reasonable security measures. Organizations not employing the GDPR regulations or not protecting the users data would face a maximum fine of around 20 million or four percent of an organizations annual global revenue. In addition, the GDPR empowers citizen to sue the organizations in the court of law that breach the GDPR. The patients health data is considered most sensitive and private data and require effective security from the insider as well as from the outsider. It is forbidden to share personal data of patients without their consent even to employees of the health organization. To fulfill the GDPR regulations and effective security of patients data we presented a framework for effective access of data based on the regulations and rights assigned to employees of the health organization.

3.2. Patient generated health data

Patient-generated health data (PGHD) is health-related data created, recorded, or gathered by patients to help address a specific health concern. PGHD include, but are not limited to health history, treatment history, biometric data, symptoms, or lifestyle choices. PGHD helps in gathering information that helps in improving care quality, reduced cost, and patients safety. PGHD are different from other types of healthcare data i.e. patient generates and collects data and later decides on whom to share the data with.

With the adoption of big data in the healthcare sector, a problematic truism is "More data equals more knowledge equal better health outcome" [42]. However, the success of PGHD in the diagnose and treatment is solely based on the willingness of the patients to share the personal information with the healthcare providers. This sharing ultimately depends on the trust and confidence of a patient in the overall sharing mechanism. With the advent of the Internet, social media, and ever-increasing means of data sharing, the whole phenomenon of healthcare data sharing will be successful only when the elements of privacy, security, trust, and ownership carefully addressed.

Omnipresent Internet and the abundance of wearable online devices have opened new horizon of healthcare data where an enormous amount of real-time data is being generated, stored, and shared on-demand by means of mobile applications, and Wifi connectivity. While PGHD provides number of opportunities, it also introduces many challenges in the healthcare industry. Under poorly designed privacy and security policies, there is a potential for exploitation of healthcare data by unlawful sharing and retrieving of sensitive information.

With ever-growing data sources, there is a need to clarify the role of PGHD in the health industry as well as in the commercial sector. For instance, National Institute of Health introduced Precision Medicine Initiative (PMI) which described as a new way of doing research that fosters open, responsible data sharing with the highest regard for participants privacy and that puts engaged participants at the center of research efforts[43, 44]. This initiative drew an assumption that in the future, healthcare data will not produce any harmful unintended consequence for the data donors. However, it is obvious that we cannot rely on these assumptions and hope that attacks on privacy will be managed and data handled with responsibility and care.

Furthermore, it must be noted that although the healthcare data related to clinical environment and PGHD are fusing, existing rules and regulations are generally only applicable to the data produced within the boundaries of the healthcare service providers. Furthermore, data produced outside the healthcare sector is not yet fully covered and affected by such regulations. Mobile technology enables users to generate large volume of real-time data. Furthermore, with improved methods of collecting, processing, and storing large amount of data being developed, the conception of the healthcare sector and mobile technologies are evolving. Consequently, it not only conveys information but also conveys medical knowledge and a mean to understand the detail and processes happening in the real time. However, mobile devices and monitoring tools generating healthcare data outside the boundaries of the healthcare sector is not covered within the rules and regulations in place to protect in-clinic healthcare data.

Furthermore, as remote sensors, monitors and devices are becoming more common and cheaper, it is becoming similar and common to use the real-time medical data to improve the wellbeing of the patients. However, data breaches, exploitation, and unlawful exposure are creating a real threat to the adoption and penetration of this trend in the healthcare services. Security and privacy are growing concerns which are creating a fear of sharing personal sensitive data. It has become obvious that these security and privacy issues must be addressed to allow a smooth penetration of PGHD in the healthcare industry.

According to [45], the social norms currently being observed in the society where people are becoming increasingly comfortable with sharing personal information and self-disclosure on the social media sites. Mobile application developers and social media sites are assuring users the safety and privacy of their sensitive data and encouraging users to share their sensitive information without the fear of exploitation. But do they really provide the security and confidentiality to the users data? With data breaches and exploitation of online data is common, it is obvious that not enough is being done and yet more efforts are required to fully secure online PGHD.

In summary, there is a dilemma in the whole phenomenon of PGHD where we need to share the data freely to fully benefit from the offerings of the real-time PGHD data but also this data must be scrutinized to make sure that only limited and necessary sharing of the data to address the privacy and security concerns.

3.3. Clinical data

Clinical data is the most important source of healthcare data. Clinical data is either gathered through the treatment of patients by the healthcare service providers or during a planned and controlled clinical trial. Clinical data can be divided into six major types:

Electronic Health Record: Electronic Health Record is by far the most abundant form of healthcare data. It is generated in the healthcare service provider facilities and usually not available to external research activities. It is a comprehensive set of healthcare data and includes a list of attributes such as personal information, age, address, lab reports, blood and urine tests, X-ray reports, insurance coverage information, medical condition, and treatment history.

Administrative data: This type mainly contains data related to admission and discharge of a patient from a healthcare facility.

Claims Data: This is mainly for the billing purposes and contains the insurance related information such as cov-

erage type, membership detail, and claims history. It can be obtained from commercial healthcare facilities or government agencies.

Patient Disease Registry: This contains information related to chronic deceases and helps understand current trends in decease such as diabetes, heart decease, cancer, HIV, and asthma. It also helps in identifying any potential outbreak of decease and also uses to manage and contain any outbreak. Patient/decease registry also helps in drawing national healthcare drives and planning health policy of a nation.

Health Survey: Health surveys analyze the overall health condition of a country and help identify most prevalent chronic diseases in a population. It results from national and local level surveys which help in research and development activities. This is one of the few healthcare data types which are primarily used for research purpose.

Clinical Trial Data: Clinical trial data refers to information collected through publicly and privately supported clinical studies from around the world. Most of the healthcare data produced and generated within the boundaries of a healthcare services provider are considered as clinical data. It is either collected during an ongoing treatment of patients or as a part of formal clinical trials. It is by far the most exclusive type of healthcare data including; administrative and demographic information, diagnosis, treatment, prescription drugs, laboratory tests, physiologic monitoring data, hospitalization, patient insurance, hospital discharge data, claims data, disease registries, health surveys, and clinical trial data [46]. As before, clinical data is of utmost importance when used to aid diagnosis, analysis and treatment of patients however digitization of the healthcare data and sharing and transmission of this sensitive data among healthcare services providers introduces the challenges of privacy and security.

Security and privacy goals are typically patients' expectations with the healthcare data and its handling. According to the Good Medical Practices, patients have the right to expect that their personal information will not be shared except for the fulfillment of the professional duties of the health services providers and that also with the consent of the patients. It is the duty of the health service providers to share the clinical data only when needed and only related information is shared and throughout this process, patient need to be aware of how much and with whom data has been shared. This is particularly important when interpreted within the newly formed regulations within GDPR.

Additionally, integrity and confidentiality of the healthcare data are also important. Healthcare system handling all clinical data must be able to protect data against any unlawful alteration and should be protected against attacks to ensure availability. It is especially challenging to detect data corruption in an electronic database due to malicious actors targeting such datasets as well as due to the untrustworthiness of the communication links used for data sharing. Guidelines and regulations such as HIPAA and GDPR do cover clinical data but the integrity of data cannot be ensured by these regulations alone. A hardened system capable of defending the healthcare data from malicious attacks and data alteration is required to ensure the integrity and availability of the information for the lawful use. Consequently, there is a need for an effective security policy to ensure authorized access to critical data.

3.4. Pharmacovigilance

Pharmacovigilance is related to data collection, detection, assessment, and monitoring for the prevention of risks and adverse effects with pharmaceutical products. Information received from patients and healthcare service providers via predefined data sharing agreements and other resources such as medical literature play a critical role in providing the data necessary for pharmacovigilance to take place. It is necessary for most countries to make this data available to the public before any new drug is introduced in the public domain. This requirement facilitates identification of the hazards associated with pharmaceutical products and to minimize the risk of potential harm to patient health and wellbeing.

The Pharmacovigilance process involves storing, sharing, and accessing of information and feedbacks by patients and health service providers to assess the possible adverse effect of drugs. Pharmacovigilance is a critical part of the healthcare industry, the outcomes of which directly affect the well-being of the general public. This process mainly involves storing, sharing, and accessing appropriate information to identify any unwanted or undesirable effect to minimize the risks. As with any other large data system involved in electronic data sharing, storing, and transferring, this phenomenon is also susceptible to confidentiality and integrity challenges.
Technological advancements and the application of Big Data in the pharmacovigilance facilitates possibility to store, transfer and share an enormous amount of data anywhere by increasing the number of patient and healthcare providers. However, the confidentiality and integrity of the data becomes an issue as with whom and what type of data is being shared. Since this data is being shared between the commercial and health services providers, the result of any data breach, an unlawful exposure or any manipulation, will result in adverse health and financial consequence [47].

4. Factors Affecting Security and Privacy in Healthcare Applications

Different healthcare service providers and data generated by them is increasingly integrated into big data technologies. This has expanded privacy and security concerns for the individuals [41, 48]. Within the big data scenario, there are several aspects of the healthcare system which need to be addressed [49]. These are briefly explained as follow.

4.1. Data access and storage

Paper-based healthcare data naturally placed a physical limit on access, edit and sharing of information. On the other hand, the electronic form of data removes these barriers and makes it virtually available to everyone. Although electronic records benefit users and healthcare providers, there can be significant adverse impact if the information is used for the purposes other than necessary health-related service. To minimize the risks associated with the electronic format of healthcare data, the following challenges require investigation.

4.2. Data Acquisition and Communication Protection

As discussed in section 2, healthcare data is diverse and multivariate collected in various formats. Therefore the data acquisition challenge within this context is twofold. Firstly, integrating different data streams to perform meaningful analytics of data is a challenge in maintaining quality of the data collected. Furthermore, data originating from different streams contains potentially diverse security contexts which require appropriate protection mechanisms.

Data ownership represents one of the major challenges in the healthcare industry. A primary question is *Do patients own their health-related data or it is the property of healthcare services providers or health insurers?* Or *it is a combined ownership?* Since data is being shared among many entities, what implications does this data sharing has over the authority and ownership when crossing organizational boundaries? It becomes challenging to ensure privacy and security of the data when it moves between different entities with different levels of authorities especially when taking into account disparate security policies.

Amount and types of data stored Healthcare data is enormous and versatile as it consists of many types such as doctors notes, lab reports, MRI, X-rays reports, and readings of vital organs. With this volume and variety of data, a challenge is how much of this data should be stored for diagnostics purposes. In such circumstances, there is a trade-off between the volume of data used and the exposure of data to security and privacy breaches as a higher volume of data can aid improved analysis and diagnosis however it also increases the risk of potential data breaches.

Storage location With the increasing volume of data, it is increasingly difficult to store all of the data at a local data store leading to a proliferation of remote data storage solutions such as data clouds. A typical characteristic of such data storage options is that they are not only located at different geographical locations but are also owned and managed by different organizations. These arrangements introduce concerns such as whether data be stored at the central location, in the cloud or at the patients premises in the case of the remote sensor. Depending on where the data is located and how it is managed, the requirements of security, access, and authority change dramatically. For example, security measures for data residing in the data center of a hospital will be different to the one in an external cloud environment which introduces a number of authentication and authorization challenges [50]

Access privileges A primary goal of the access control as a measure to achieve security and privacy for healthcare data is that a person only has access to the required data which they are entitled to. Within this context, users can be categorized into several classes to ensure legitimate access to crucial information. Some users such as doctors and nurses may have a read and write access to the patients data they are attending to while others such as health insurers might only need read access to their clients data. In any case, the principle of least privilege should be followed when assigning privileges to different user groups.

Patients consent As a significant part of healthcare data is related to patients i.e. clinical, personal and diagnostic data. This introduces a need to engage individuals who are represented by the data items and therefore to seek their consent before sharing and processing data. Recent advancements such as the GDPR have brought special emphasis on personally identifiable data and the need to have consent from individuals which can impact a security policy and respective protection mechanisms. However, there can be situations such as emergency cases where data of a patient needs to be shared with a party previously not authorized for sharing. Such circumstances mandate special attention and distinct efforts to address them.

4.3. Data analytics

Large-scale healthcare data is required to be analyzed to extract valuable information. Currently, most of the data is paper-based but with time it set to be digitized. Healthcare data analytics are tools used to perform analysis on the raw healthcare data to detect patterns and trends. This helps in improved diagnosis as well as helping with measuring the effectiveness and progress of an ongoing treatment. Clearly, it will help in a higher quality of the healthcare services and reduced cost. Premier, U.S. Healthcare Alliance Network has reported that the use of healthcare analytics helped them save an estimated 29,000 lives and reduction of approximately 7 billion in healthcare cost [51]. Obviously, access to healthcare data will have risks. Privacy and security would be a concern accessing these records as a query might generate unwanted data or sensitive and secretive data. Safeguards such as an up to date antivirus, setting up a firewall, encryption of sensitive data and multifactor authentication are required to protect confidential healthcare data [52]. Auditing mechanisms are needed to enable the flow of data yet securing against any unlawful access. Data anonymization provides an effective option for privacy however it limits conducting rigorous analytics of the healthcare data.

4.4. Heterogeneous regulations

Healthcare data is diverse in nature and usually scattered around different healthcare entities. In the case of

big data, this data can be distributed theoretically anywhere in the globe. Different entities are regulated by different rules and transfer of data between different entities with different regulations can be tricky. For example, patient's consent for disclosure of data may be compulsory for some while it is optional for others. When data moves with health care service providers falling under different rules and regulations, privacy and security may become an issue even through legal disclosure of data.

5. Security and Privacy Challenges in Healthcare

Healthcare systems collaborate in nature. This is because it has number of stakeholders such as physicians, nurses, lab technicians, and pathologists working together to maximize the effective use of raw and defined healthcare data. Each stakeholder generates heterogeneous data such as physical exams, lab reports, diagnostics notes, clinical notes, imaging analysis, patients observations and interviews, and progress and outcomes of the of therapies and treatments. The use of latest and emerging technologies, sensors and sharing techniques, and information and communication techniques make it easier to generate and share healthcare data but also introduces challenges with respect to security and privacy. This is because electronic healthcare data is susceptible to unlawful access, compromise integrity, and unauthorized distribution [53, 54, 55].

Another major challenge with big data applications in the healthcare sector is complex and distributed nature of healthcare data, diverse schema and standards, and rapid growth of new health terminologies and ontologies. The major challenge here is not the lack of data but the lack of information to support decision making, planning, and strategy. Following are some of the issues with the adoption of big data in the healthcare [46]:

5.1. Resistance to change

Healthcare system is traditionally lagging in adoption of the new technologies compared to the other sectors such as banking or oil industry. This is because of lack of the technical administrative support, legislative issues, lack of trust and slow in changes to the medical practices and lack of expertise in the ICT.

5.2. Fragmentation of the healthcare data

Healthcare data is fragmented and distributed across the healthcare system consisting of legacy as well as modern equipment with limited interoperability capabilities. It is, therefore, challenging to integrate different data types due to different schema, formats, and standards.

5.3. Ethical Challenges

Access to complete and comprehensive data related to a patient in a timely manner is vital to effective diagnosis and treatment. However, sharing of information among many stakeholders become difficult due to ethical issues including confidentiality and integrity of the patients data, control, and extent of access to medical record, ownership, and governance of the healthcare data and commercialization of the healthcare data.

5.4. Proliferation of healthcare standards

Standards are agreed upon specifications that allow different systems, tools, and platforms to work with each other. Healthcare sector lacks a central single standard for different sources of health data. Due to different layouts and formats of diagnostic reports, examinations, drugs, and decreases, it is challenging to integrate data from different sources and stakeholders into a single entity.

5.5. Rules, laws, and regulatory bodies

Other factors that must be considered while using big data analytics for healthcare are rules and laws. Healthcare privacy and security not only concern with the expectations but also to norms involving professional practice, privileges, protected communication, and duty of confidentiality, as well as to data collection, distribution, and retention [56]. In many developed countries, different health acts not only provide a guidance in handling healthcare data but also set benchmarks on setting guidelines on how to whom and to what extent of healthcare data be enclosed.

5.6. Technological challenges

In todays business environment, a multitude of devices, users, and enormous traffic all combine to create a proliferation of data. Security and privacy issues are the center for the optimum use of the big data. Traditional tools to handle such data is not sufficient and new tools and methods are required to essentially handle the large volume of healthcare data. It can be concluded that the list of issues mentioned above, privacy and security concerns are vital to the penetration and adoption of big data deployment in the healthcare sector. Trust in privacy and security of the healthcare data results in the level of information a person is willing to share [56]. It is important to discuss different types of healthcare data and their unique privacy and security issues to fully grasp the nature of the problem we currently face in generating, storing, retrieving, and processing of the healthcare data.

6. State-of-the-art for Security and Privacy in Healthcare Systems

Data access, storage, and analysis are not unique to healthcare system. Any large volume of data with many stakeholders with different interests and authorities constitutes similar security and privacy challenges whether online shopping, financial services, defense secrets or even public service providers. Any measure introduced to protect the sensitive data can be applied in all these sectors

regardless of the nature of the data or institution. It is an evolutionary process where many security and privacy measure are already in place, being updated to improve the security or to counter the newly discovered loopholes.

6.1. Access Control

The most challenging aspect of a large healthcare data network is the security administration. Healthcare system is a complex system with diverse data and the ever-increasing number of users with different requirements for types of data as well as durations. There are many access control models but Mandatory Access control (MAC), Discretionary Access control (DAC) and Role-Based Access control (RBAC) are the most popular [57]. MAC model gives total control to the security policy administrator and the user has no control or authority to override established policies. Security policy administrator defines the usage and resources and their access policy and this policy defines who has access to which files. This model is suitable where confidentiality is the main concern such as defense and national security matters. DAC model gives the authority and control to the end users where end users have full control over the resources they own and owners decide who can access their resources. Finally, the RBAC model is based on rights and access to resources according to the membership to predefined groups. RBAC helps in making a mechanism where data is accessed on the need to know basis. It results in less complexity of the system, reduced cost as well as protection of the sensitive data beyond its required and necessary disclosure [49].

As mentioned earlier, access control of the resources and information Big data offers is a key feature of the overall security and privacy aspects of healthcare. As we know, healthcare is a dynamic sector where the role is not defined solely by the profession or position of a user in the organization but also depends on the situation [57]. In view of the significance and relevance of the above-defined access control models in the healthcare sector, MAC is one of the strictest and most secure. It is a type of access control model where confidentiality is the main priority and where all control and authority rest with a central administration. It suits military and defense sectors better as the main priority is to protect the assets from unlawful access. It is static in nature and the design requires a lot of planning before the implementation. Even after implementation, it requires a huge system management overhead to add new objects, add new users, or modify the rights of existing users. These aspects make this model unsuitable for the healthcare sector due to its dynamic nature.

DAC model provides the flexibility needed by the healthcare sector as it rests the authority with the users but also make it more prone to security and privacy risks as users have full control on the assets they created and a user can set who can access their data. Again, it is obvious that risks associated with this model make it unsuitable for the healthcare data. RBAC is currently most popular mode of access control in the healthcare sector as it defines the privileges and rights of a user based on its membership to a group or groups based on their function in the healthcare sector. From a healthcare point of view, it means groups like GP, nurses, lab technicians or X-rays technicians. But again, lately, some of the drawbacks of this model in the healthcare sector have emerged as there is no way to provide individual users with the rights over and above the privileges assigned to the group they belong to. Since healthcare staff goes through many different situations, their demand for healthcare information can change abruptly. One of the most important situations in the healthcare sector is emergency. In the face of an emergency, the roles of the user and the demand for data changes.

More recently, Attributes Based access control (ABAC) found its way into the healthcare sector as it provides the flexibility needed to handle the situations in the healthcare sector where responsibilities and demand for data change. ABAC is based on users attributes as well as on resources, object, and environmental attributes. It works on the Boolean logic where access to data based on IF and Then statements [58]. In ABAC, permission to access objects like files, images, and reports are not simply based on the subject but depends on the attributes of the subject. Here subject means any user who is accessing the healthcare data. These attributes are static like name, position and role and dynamics like role, situation, environment, and location. Based on edit and entry, healthcare data can be generalized into two types: *Static Data* usually stays the same in the normal scenario and some of the types are patient personal data like name, sex, blood group, allergies, past medical treatment and records and insurance information. On the other hand, *Dynamics Data*, such as care plan, progress notes, readings of the vital organs, or medical reports, usually modify, update, or delete on the regular basis.

In the healthcare system, diversity is found not only in the data but also in the user and their level of access to the data. The access depends on other stakeholders consent and input and it creates a complex web where an entity has some privileges and do other tasks after getting the approval from suitable entities within a healthcare system. To understand the complexity of the system, we analyze the mechanism of accessing data from a single patient [59]:

Administration has the widest access to the healthcare data compared to other stakeholders and they are mainly responsible for controlling the types of access for the other stakeholders. This data includes medical, personal, and financial information. But again, this access is restricted. They can add past medical records and can make medical entries but in the latter case, requires the consent and approval of a doctor. They are also the only entity which can authorize to delete medical data but must abide by the laws on data retention period.

Doctors usually have full access to the medical records of their patients and can add medical entries and private notes. Private notes are not visible to the other health

workers and administration and only shared between doctors and their patients. In case of an emergency or a visiting doctor attending other doctors, the patient gets the temporary access to the medical record by the consent of the patient and notification to the administration.

Healthcare workers are required to sign the confidentiality agreement before gaining any access to healthcare data. Normally they have the access to the care plan and can add progress notes. As far as emergency data is a concern, crucial data is usually available to all healthcare workers so they have all the information require to handle an emergency. It must be noted that they normally have access to recent medical records and historical records are usually not available to them but can be obtained by requesting administration. Auditing is implemented by having a logging mechanism where any unlawful attempt is recorded and if a pattern is discovered, concerned authority is informed for suitable and preventive action.

Patients have the right to access their own data at any time which also includes the private notes of doctors. Of course, this access to be only read as they shouldnt be allowed to alter the data.

Other users such as volunteer helpers, visiting physiotherapist, social workers, and community service members also require some sort of healthcare data that must be authorized through administration with the time stamp and in the end with the consent of the patients. It is obvious from the above discussion of different stakeholders and their requirements for the access to the healthcare data that RBAC model is most suitable as the roles and requirements are ever evolving and the mechanism of accessing healthcare information needs regular reviews and appropriate amendments [60].

6.2. Encryption

Encryption is used to ensure the security and provides protection against eavesdropping and skimming. There are many types of encryption are currently deployed in the healthcare data, some work on the hardware level and other work on the software level. For best results, it is necessary to deploy encryption on both software and hardware levels. Some encryption is based on symmetric key while other asymmetric key based. Again, as there are many solutions available for encryption and are being deployed in the healthcare sector, each has some limitations.

In 2009, IBM developed a homomorphic encryption technique which allows processing of the data without decryption. It gives a huge boost to the processing time as well as the preservation of crucial resources. But since it is based on an algebraic algorithm, a different ciphertext can be created from a plaintext. Also, homomorphic encryption is not semantically secure and cannot provides any means for verification [61]. Another encryption technique used is attribute-based encryption [62]. This technique called for an encryption of data before being shared with others. Basically, the owner of the data has the authority to encrypt the data and allows which entities will have the access to the data and more precisely which part of data. Owner will share the keys for decryption with the selected entities which later will use these keys to decrypt data. Owner of data only retains the privilege to revoke or grant the access.

The main issue with any encryption technique is key management. Not all users are experienced with keys to managing their own keys and on the other hand, a central key authority will be easily overwhelmed by the number of keys they must manage in the case of the healthcare sector.

6.3. Authentication

Authentication techniques are used to verify the source of data to make sure data is coming from the source it claims to be. There are many authentication techniques used in the healthcare sector like password and digital signature but can be seen that the most common form of authentication method used in the healthcare sector is an identifier with a password [63]. There is much software available to retrieve the passwords and users are usually choose any easier password for convenience. Also, it is common for the users to expose their passwords through social engineering and carelessness. One common habit is to write down the password on a piece of paper and keep it in an easily accessible place. A better authentication system is to have a credential system where only those already possess the legitimate credential can access the system. In a credential system, a user obtains a credential from an organization and later display possession. A user can perform some cryptographic operation on the credential by digital signing. It is more secure than password as it cannot be retrieved by guessing.

6.4. Policy development

As mentioned earlier, the healthcare sector is undergoing a revolution and an ever-increasing amount of data is being generated, share, and store. And mentioned the complexity of the overall system, where who access what data is changing over time. A set of fixed policy currently exists in many healthcare services are inadequate and a threat to the security and privacy. There is a need for a dynamic and scalable policy on the data access and sharing. There must be a mechanism where the rights of all stakeholders to data need to be revised based on their needs without compromising the privacy and security. Also, there must be an authority who can monitor and later revise the access to the data with the consent of the patients. Most importantly, there is a need to bring awareness about the importance of the personal healthcare information and consequences of any unlawful access or alteration of data. In the end, it must be obvious that the mechanism of revising the rights to access type and extent of data cannot be done manually and an automation is required where conditions are set for a change to happen automatically.

6.5. Data anonymization

Data mining is a process of analyzing data to identify the pattern and to extract information from a large amount of data and thus present a serious security and privacy issue. As mentioned earlier, healthcare data is enormous, heterogeneous and distributed in nature. There is a need to standardize this data for a better analysis and extraction of useful information but any effort in this direction also ensures data exploitation unless proper security and privacy measures are in place [49]. Data anonymization is in place to hide some attributes of a patient like a name, age, gender, and address. But again, this anonymization is not very effective as still any unlawful extraction and analysis of the data will lead to some sort of pattern and will provide sensitive information.

Anonymization is a complex phenomenon where anonymization is not an issue but later the extraction of the information from the anonymized data is a challenge. Due to the diversity of the healthcare data and different stakeholders demanding access to different part of the whole data, there arises a question of what level of anonymization is suitable for all users. It is understood that data need to be anonymized before it being shared but what information needed by entities like doctors and insurance companies is different and similarly the level of data mining capabilities required to extract information is diverse. Besides, while ethical practices are well defined for the primary users like doctors, patients and nurses, there is still a lot of work need to be done for the vast array of disclosures to secondary users like insurance companies and health care evaluators [49].

7. Comparative Analysis of Existing Privacy Assurance Models for e-Healthcare

Privacy assurance within e-healthcare can be studied from various dimensions highlighting specific privacy requirements. In this paper, we focus on Electronic Transfer of Prescription (ETP) and access for emergency services dimensions to conduct an in-depth study of the challenges for privacy assurance within e-healthcare in general and ETP in particular.

7.1. Electronic Transfer of Prescription

Electronic Transfer of Prescription is a generation, transmission, and processing of medical prescription performed electronically instead of the traditional paper-based system. It allows doctors, physicians, and pharmacists to transmit and share error-free, understandable and accurate prescriptions. A prescription is typically originated at the doctor office and destined for the pharmacist. ETP is therefore envisaged to reduce the risks involved in manual paper-based prescriptions which serves as the primary motivation for its adoption in the healthcare sector. Compared with traditional paper-based approaches for a prescription, ETP has a number of advantages including; improved patient safety and care through error-free sharing of prescriptions as well as speeding up the process of executing a prescription. Furthermore, ETP also enables reducing the overall cost by providing access to less expensive drug alternatives and avoids duplicate prescriptions. Similar to a generic electronic solution, As with any electronic form, ETP is also susceptible to security requirements such as those translated into confidentiality, integrity, and availability.

The UK government has initiated implementation of ETP in the NHS to handle more than 500 million prescriptions annually and a central Prescription pricing authority (PPA) has been established for processing all prescriptions and thus is a central part to any transaction in an ETP system [64]. With regards to addressing security and privacy challenges, three ETP models have been considered which include; 1) Transcript Consortium Model – where a prescriber generates a prescription, digitally signs it and generates a barcode. Prescriber also sends an encrypted electronic prescription to the PPA as per requirement. Patient than takes the barcode to any pharmacy part of ETA and pharmacist uses the barcode to validate the digitally signed prescription and dispensed drug. 2) Pharmacy 2U Consortium Model – is based on the direct communication between the prescribers and pharmacies where the patient is asked for the choice of pharmacy for pick up. GP than digitally signed the prescription and sends it directly to the selected pharmacy. A copy of the same prescription signed by the GP also sent to the PPA. 3) SchlumbergerSema Consortium Model – is based on a relay system called fexiscript as an intermediate data store has introduced between the GPs and pharmacists. GPs dont have any direct communication with the pharmacists and any communication from GP to pharmacist must be handled by a data store an intermediate entity. GP digitally signed a prescription and encrypt it for the data store and send it to the intermediate data store.

In addition to these models, Salford Model [29] was developed to address the shortcomings of the earlier models. Like Flexiscript, Salford model is a relay based system which in this case is a prescription store. In a typical scenario, a prescriber generates a prescription, digitally sign it and symmetrically encrypt and sends it to prescription store. This message also contains an encrypted symmetric key for the PPA. It must be noted that in this model there is no direct communication between GPs and pharmacists with PPA and only the prescription store has a communication channel with PPA. The patient is provided a paper with reference barcode for accessing the prescription and asymmetric key barcode to decrypt the prescription in the prescription store. These two barcodes help a pharmacy to retrieve and decrypt the prescription from the prescription store. Pharmacist, upon dispensation of prescription, sends a dispense message encrypted for PPA to the prescription store. Prescription stored periodically retrieves information about dispensed and non-dispensed prescription from the prescription store. In the case where the patient has a preferred pharmacy, GP sends an email to the

chosen pharmacy which again contains reference barcode and symmetric key barcode and in this case, the prescription can be made ready for pick up by the pharmacy prior to patients arrival at the pharmacy.

7.2. Analysis of existing ETP privacy assurance models

In view of the security and privacy requirements of the above-defined scenarios, a rigorous analysis of existing privacy assurance models for confidentiality, availability, and integrity is paramount. Such comparative analysis is envisaged to facilitate enhancements aimed at improving the state of the art to achieve specific security and privacy requirements. Within this context, a detailed comparison of different ETP models with a summary of comparison is presented in Table 1

As is evident from the table, the Salford model attempts to address all security requirements for ETP including freedom in choice of pharmacy and non-repudiation. However, the comparison has also highlighted two limitations in Salford mode i.e.

- Salford model proposes to X.509 & PERMIS to achieve authorization which is focused at asserting the identity of the prescription issuing authority. However, it does not mitigate against malicious attacks such as masquerading with respect to the use of a prescription.
- The Salford model attempts to address duplication and fraudulent use of prescription by using encrypted prescription where a prescription can only be used once. However, this mechanism is limited in protection against masquerading and ID theft attacks where a prescription is misused by an imposter.

The above limitations highlight a gap in existing privacy assurance models for e-healthcare in general and ETP in particular. We address these limitations as part of our enhanced privacy assurance model in the next section.

8. Enhanced Privacy Assurance Model for ETP

In this section we present our reference model for securing the healthcare data when data is accessed by multiple medical paractioners.

8.1. Usage Scenario

We are adopting the same usage scenario mentioned in [65]. A patient, named Gorge, is recently diagnosed with a gastric cancer. For many patients, chemotherapy and radiation therapy after stomach surgery increase the chances of cure. For the treatment and surgery, Gorge entered a recommended cancer-treatment center. George also has a general family practitioner whom he regularly visits for his treatment and medical consultation. Upon entering the hospital, Gorge also has an attending doctor from the hospital. On the analysis and during the initial treatment Georges health condition deteriorate and has cause some complications that his attending medical paractioners doctor would like to have an expert opinion and consultation for Gorges treatment from different medical specialist doctors, including Gorges specific general practitioner because he is fully informed about Gorge s medical history. The invited practitioners are specialized in different subjects and all the involve person requires to have the medical records for the analysis with the request based on the HIPAA minimal disclosure principle. Furthermore, the consultation result, such as the diagnosis and treatment suggestions, should be signed and certified by this group of specialists and practitioners. The medical certificate with their signatures is sent to Gorge. If Gorge would like to share this medical information with her loved ones and her family physician, he can put the new medical certificate into her PHR database.
In this setup the paractioners and the Gorge requires protection of their private data. From the practitioners point of view he need to know how to securely obtain the medical history of Gorge and how to ensure that the received history is obtained after the consent of Gorge. This relates to the problem of secure authorization of the medical reports. Similarly, for patient, Gorge needs to be ensured the allowed person could only have secure access to his medical records. Our proposed model ensure the privacy and security of medical data for both paractioners as well as patient.

8.2. Proposed Model for ETP

Through the comparison performed in section 7, it is highlighted that Salford Model satisfies major security requirements for ETP concerning confidentiality, authorization, integrity, and availability. However, there are limitations for Salford model specifically with respect to authenticated access to prescription and with respect to data transmission through barcodes. For instance, although encrypted prescriptions address the confidentiality requirement, these do not mitigate against scenarios where a prescription is lost by the patient or misused by someone other than the patient (impersonation and ID theft). Furthermore, the data that can be communicated using barcodes are limited in type and volume and therefore has seen them be replaced by QR codes.

Within this context, we present specific enhancements to the Salford model targeting mitigation against the above limitations by incorporating multi-factor authentication and the use of QR codes. The Enhanced Salford Model is presented in Figure 1. The important features of enhanced Salford model are as follow:

1. Replacement of Barcode with QR code to improve customization, storage ability, and error correction.
2. SMS notification to the patient whenever a pharmacy tries to access prescription from the prescription store using QR code and request for the verification code. This will achieve multi-factor authenti-

	Transcript	Pharmacy2U	Flexiscript	Salford
Authentication	No	No	No	X.509 & PERMS
Encryption	No	Yes	Yes, but data store can decrypt	Yes
Confidentiality	Yes	Yes	Weak	Yes
Freedom in choice of pharmacy	Yes	No	Yes	Yes
Digital signature	Yes	Yes	Yes	Yes
Duplicate/fraudulent prescription	Weak	Weak	Weak	Weak
Availability	Yes	No	Yes(revert to paper based)	Yes

Table 1: Comparison of ETP Models

cation mitigating threats such as masquerading and ID theft.

3. SMS notification to the patient when PPA try to access prescription from the prescription store using QR code and request for the verification code. This will facilitate multi-factor authentication mitigating threats such as masquerading and ID theft.

QR code is 2-dimensional barcode consists of black squares in a grid form on a white background. It can be read by an imaging device like a camera and processed using Reed-Solomon error correction until the image is interpreted appropriately. The obvious difference between QR code and the barcode is the way how they store data. While barcode stores data vertically, QR code can store data vertically as well as horizontally. QR code has many advantages over barcodes. Some of the advantages are as below:

- QR code is much smaller in size as compare to barcode
- QR code can store 100 times more information than a barcode
- QR code is easier to use and easier to read as it can be scan over 360 degrees thus eliminating and interference and negative effects from background code offers greater error margin as compared to barcode
- QR code can store informations in term of characters, symbols, text, and control codes

As shown in Fig 1, a patient is asked to select a personal code when his account is being set up in the healthcare system along with other information like address and mobile number. When a prescription in the form of a QR code is handed over to the patient, or scanned by the patient or sent digitally to the patient by GP, the patient can take it to a pharmacy of choice. When pharmacist scans the code to retrieve the prescription, a notification is sent to the patient's registered mobile. Patient must authorize the access by using his personal code set during account setup. Only upon confirmation by the patient, the pharmacist can retrieve the prescription from the prescription store. Later when the transaction is completed, another is SMS sent to the patient to acknowledge the completion of the transaction and again the patient must confirm using personal code. The same procedure will follow when GP send QR code directly to the pharmacy. The same process will follow in the case of home delivery whereupon delivery patient will receive an SMS notification and again must respond by using his personal code to acknowledge delivery. This two-factor authorization will ensure additional security in the model overall as the patient will be total control of which pharmacy retrieve prescription and to avoid fraudulent claims. Even when PPA will try to access a patient prescription, the patient will receive an SMS notification as demonstrated in Fig 1.

9. Recommendations and open challenges

Proliferation of contemporary and emerging technologies has defined our lives. From traditional Internet based systems to emerging *smart* technologies, technology has influenced healthcare aiming to achieve benefits such as real-time monitoring, diagnosis, and treatment from virtually anywhere globally. In this section, we present a discussion of the recommendations and open challenges to achieve privacy assurance within smart healthcare.

Security of mobile healthcare platforms Mobile technologies and the Internet has made it possible to access medical data, images, and remote monitoring anytime anywhere. This phenomenon has transformed the healthcare sector into a dynamic environment where virtually everything is available anywhere. However, use of the Internet and mobile technology introduces risks of privacy, integrity, and confidentiality as a distributed structure with potentially unlimited users poses a great risk of making data exposed beyond required limits. A mobile healthcare system helps improve patient care, quality of healthcare services, professionalism and productivity and reduction in cost. It is extremely significant that patients and healthcare workers have confidence in the confidentiality, integrity, and security of the mobile technology for an effective deployment of the same in the healthcare sector. Currently, a number of measures are available to address the privacy and security issues of the mobile technology in

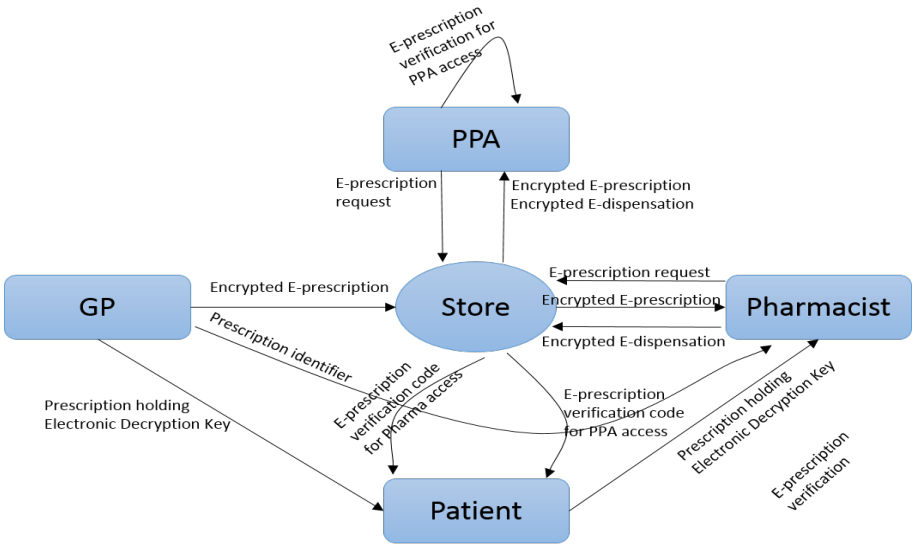

Figure 1: Enhanced Salford Model

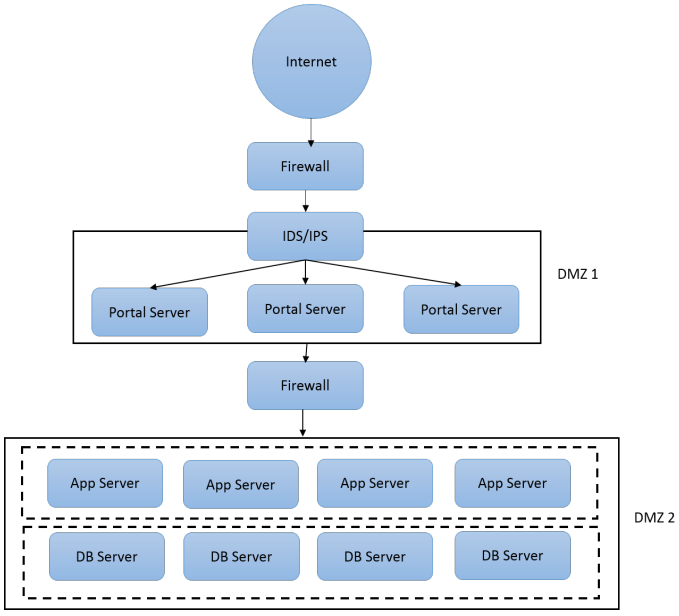

Figure 2: Secure architecture for healthcare database

the healthcare services such as data encryption, database security, security policies and defining extent of medical data sharing [30].

Security of mobile devices: Users not only needed to be authenticated to access the healthcare data from their mobile devices but additionally, this access should be from mobile applications and this mobile application should be designed in such a way to force the users to have some sort of security features on their mobile devices like password or biometrics. This way healthcare-related mobile application would not run on an open mobile device without any security feature enabled. Multifactor authentication must be enabled to enhance security from credential theft. Additionally, there should be a two-factor authentication where an attempt to access a medical record automatically sends a message to the patient mobile device and only upon approval by the patient, medical record can be accessed. This way a patient can control the access to the medical record as well as becomes aware of any attempt to access its personal information.

Multi-factor authentication: Multi-factor authentication facilitates protection against masquerading and ID theft attacks by incorporating additional user input in the authentication process. Whilst multi-factor authentications, access upon acknowledgment by patients, and ABAC can be applied to improve the privacy and security measures in the healthcare sector, there are still areas which require further improvement. It is recommended to use Security Information and Events Management (SIEM) software to analyze real-time analysis of the security alerts as well as investigating the source of attacks. Further auditing of the database is required to make sure that legitimate users are not exploiting their privileges. Here is a scenario to analyze the security and privacy issues of accessing healthcare data using mobile technologies and their related remedies.

Medium: Since in the case of mobile technologies, the medium used to transfer information is wireless, anyone can have access to this information in the air. It is of utmost importance that strong encryption is in place to make sure that even in the advent of an unlawful capture of this information; a hacker would not be able to extract the sensitive information.

Security of healthcare database: Servers not only store the data but also make it possible to do data mining

using analytical tools for understanding the data. Since database servers are a source of sensitive and confidential data, it should be secured through authentication [66, 67], applying privacy enhancing tools (e.g differential privacy, anonymization) [68, 69] role-based access control [70], security policies to restrict the access [71] and using blockchain technology for the healthcare 8327543. A comprehensive architecture to secure healthcare database is presented in a Figure 2.

Dynamic security policy The healthcare sector is dynamic where the role of a user changes with the location and situation. For example, a physician requires access to a patient data in an emergency scenario which under normal situation is denied. Although RBAC used intensively in the healthcare sector, it has limitations when it comes to allowing access to resources based on the location, situation, and time considerations. Attribute-based access control has recently been deployed in the healthcare sector to address the limitation of the RBAC as accesses control is not fixed but based on different attributes which creates an *if than* scenarios and access is based on fulfilling certain conditions and access to resources becomes flexible and scalable. Therefore, there is a need for a comprehensive security policy with clear and well-defined attributes as well as defining the limits on what data should be available in a certain situation. This policy should also include auditing features where any breach of data or attempt for an unlawful access should be detected and recorded for appropriate preventive actions.

Standardization of healthcare data: Healthcare data consists of structured and unstructured data, images, handwritten notes, and graphs. Since this data shared between different entities like healthcare centers and insurance companies, there is a need to have a standardized format so sharing, flow, and understanding of data can be achieved in a seamless manner achieving integrated operation across heterogeneous systems.

Patients awareness: Since patients are considered the primary owner of their healthcare data, it is important that they must be aware of the risks and benefits of their confidential information sharing. It is common for a patient to expose their data to the unrelated person leading to masquerading and ID theft attacks. Within this context, patient awareness can be raised with respect to risks and benefits of sharing their personal data. They should be aware of whom to share the data and up to what extent and under what conditions.

10. Conclusion

Internet of Things have revolutionized diverse domains including healthcare with emerging applications to achieve improved well being and overall quality of life. Context-awareness and mobility afforded by IoT has a profound role in this. However, healthcare data is typically sensitive demanding specific measures to achieve in-depth protection against its misuse. Although the use of smart devices opens new horizons for emerging smart healthcare paradigm, it also introduces new challenges with respect to the privacy of the data involved and the context-aware mobility of smart healthcare applications. This paper has therefore focused on privacy preservation within smart healthcare with a special emphasis on privacy assurance challenges within the Electronic Transfer of Prescription (ETP). It mainly emphasized the role of the patient in any privacy and security measure and the recently enacted GDPR law has emphasized this further. Based on rigorous analysis of existing privacy preservation models we propose an enhancement for the widely used Salford model to achieve privacy preservation against masquerading and impersonation threats. We conclude that any security and privacy measure cannot be effective in the absence of an effective and scalable security policy. A flexible and automated security policy to accommodate the ever-changing environment, roles and responsibilities of the healthcare workers can assure a secure healthcare system capable of protecting a patients confidential and sensitive information.

References

[1] R. S. Shankari Bhattacharyya and A. Thangavelu, "Context aware health care application," *International Journal of Advancements in Technology*, vol. 2, pp. 461 – 470, 2011. [Online]. Available: https://www.omicsonline.org/open-access/context-aware-health-care-application-0976-4860-2-461-470.pdf

[2] H. Kaur, N. Kumar, and S. Batra, "An efficient multi-party scheme for privacy preserving collaborative filtering for healthcare recommender system," *Future Generation Computer Systems*, vol. 86, pp. 297 – 307, 2018.

[3] B. Thuraisingham, M. Kantarcioglu, E. Bertino, J. Z. Bakdash, and M. Fernandez, "Towards a privacy-aware qunatified self data management framework," in *Proceedings of the 23Nd ACM on Symposium on Access Control Models and Technologies*, ser. SACMAT '18, 2018, pp. 173–184.

[4] O. Hamdi, M. A. Chalouf, D. Ouattara, and F. Krief, "ehealth: Survey on research projects, comparative study of telemonitoring architectures and main issues," *Journal of Network and Computer Applications*, vol. 46, pp. 100 – 112, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804514001672

[5] "The influence and use of individual patient comments — agency for healthcare research & quality." 2017. [Online]. Available: https://goo.gl/NUWeaG

[6] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big data: Issues and challenges moving forward," in *2013 46th Hawaii International Conference on System Sciences*, Jan 2013, pp. 995–1004.

[7] H. Chen, R. H. L. Chiang, and V. C. Storey, "Business intelligence and analytics: From big data to big impact," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012. [Online]. Available: http://www.jstor.org/stable/41703503

[8] D. He, R. Ye, S. Chan, M. Guizani, and Y. Xu, "Privacy in the internet of things for smart healthcare," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 38–44, APRIL 2018.

[9] M. A. Sahi, H. Abbas, K. Saleem, X. Yang, A. Derhab, M. A. Orgun, W. Iqbal, I. Rashid, and A. Yaseen, "Privacy preservation in e-healthcare environments: State of the art and future directions," *IEEE Access*, vol. 6, pp. 464–478, 2018.

[10] M. Al Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of Medical Systems*, vol. 36, no. 1, pp. 93–101, Feb 2012.

[11] A. Bagula, M. Mandava, and H. Bagula, "A framework for healthcare support in the rural and low income areas of the developing world," *Journal of Network and Computer Applications*, vol. 120, pp. 17 – 29, 2018.

[12] "Gdpr portal: Site overview," 2018. [Online]. Available: https://www.eugdpr.org/

[13] N. Mohammed, B. C. Fung, P. C. Hung, and C.-k. Lee, "Anonymizing healthcare data: A case study on the blood transfusion service," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '09, 2009, pp. 1285–1294.

[14] G. Cormode and D. Srivastava, "Anonymized data: Generation, models, usage," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '09, 2009, pp. 1015–1018.

[15] F. K. Dankar and K. El Emam, "The application of differential privacy to health data," in *Proceedings of the 2012 Joint EDBT/ICDT Workshops*, ser. EDBT-ICDT '12, 2012, pp. 158–166.

[16] C. Esposito, A. D. Santis, G. Tortora, H. Chang, and K. K. R. Choo, "Blockchain: A panacea for healthcare cloud-based data security and privacy?" *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31–37, Jan 2018.

[17] A. Iyengar, A. Kundu, and G. Pallis, "Healthcare informatics and privacy," *IEEE Internet Computing*, vol. 22, no. 2, pp. 29–31, Mar 2018.

[18] H. A. A. Hamid, S. M. M. Rahman, M. S. Hossain, A. Almogren, and A. Alamri, "A security model for preserving the privacy of medical big data in a healthcare cloud using a fog computing facility with pairing-based cryptography," *IEEE Access*, vol. 5, pp. 22 313–22 328, 2017.

[19] C. S. Kruse, B. Smith, H. Vanderlinden, and A. Nealand, "Security techniques for the electronic health records," *Journal of Medical Systems*, vol. 41, no. 8, p. 127, Jul 2017.

[20] R. J. Anderson, "A security policy model for clinical information systems," in *Proceedings 1996 IEEE Symposium on Security and Privacy*, May 1996, pp. 30–43.

[21] K. Abouelmehdi, A. Beni-Hessane, and H. Khaloufi, "Big healthcare data: preserving security and privacy," *Journal of Big Data*, vol. 5, no. 1, p. 1, Jan 2018.

[22] D. Charles, G. Meghan, and F. F. Michael, "Adoption of electronic health record systems among us non-federal acute care hospitals," in *ONC data brief 9 2008-2013*.

[23] S. Krawczyk and A. K. Jain, "Securing electronic medical records using biometric authentication," in *Audio- and Video-Based Biometric Person Authentication*, T. Kanade, A. Jain, and N. K. Ratha, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1110–1119.

[24] A. K. Das and A. Goswami, "An enhanced biometric authentication scheme for telecare medicine information systems with nonce using chaotic hash function," *Journal of Medical Systems*, vol. 38, no. 6, p. 27, Jun 2014.

[25] E. Ball, D. W. Chadwick, and D. Mundy, "Patient privacy in electronic prescription transfer," *IEEE Security Privacy*, vol. 99, no. 2, pp. 77–80, March 2003.

[26] S. Mukherjee, I. Ray, I. Ray, H. Shirazi, T. Ong, and M. G. Kahn, "Attribute based access control for healthcare resources," in *Proceedings of the 2Nd ACM Workshop on Attribute-Based Access Control*, ser. ABAC '17, 2017, pp. 29–40.

[27] H. S. G. Pussewalage and V. A. Oleshchuk, "An attribute based access control scheme for secure sharing of electronic health records," in *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Sept 2016, pp. 1–6.

[28] S. Alshehri, S. P. Radziszowski, and R. K. Raj, "Secure access for healthcare data in the cloud using ciphertext-policy attribute-based encryption," in *2012 IEEE 28th International Conference on Data Engineering Workshops(ICDEW)*, vol. 00, 04 2012, pp. 143–146.

[29] D. Chadwick, "The secure electronic transfer of prescriptions." [Online]. Available: https://goo.gl/7ENZGF

[30] N. Greta, C. Maria, and G. Claudia, "The role of mobile technologies in health care processes: The case of cancer supportive care," in *J Med Internet Res*, 2015.

[31] T. Li, Z. Huang, P. Li, Z. Liu, and C. Jia, "Outsourced privacy-preserving classification service over encrypted data," *Journal of Network and Computer Applications*, vol. 106, pp. 100 – 110, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804517304356

[32] "An introduction to health care administrative data," 2017. [Online]. Available: https://goo.gl/DPtGHA

[33] "Documentation how important is it - crozer-keystone health system," 2017. [Online]. Available: https://goo.gl/9FkVpg

[34] H. K. Patil and R. Seshadri, "Big data security and privacy issues in healthcare," in *2014 IEEE International Congress on Big Data*, June 2014, pp. 762–765.

[35] N. Mohammed, S. Barouti, D. Alhadidi, and R. Chen, "Secure and private management of healthcare databases for data mining," in *2015 IEEE 28th International Symposium on Computer-Based Medical Systems*, June 2015, pp. 191–196.

[36] "Nhs surveys: Focused on patients' experience: Improving healthcare," 2017. [Online]. Available: https://goo.gl/rnwhkw

[37] L. Wu, L. Yuan, and J. You, "Survey of large-scale data management systems for big data applications," *Journal of Computer Science and Technology*, vol. 30, no. 1, pp. 163–183, Jan 2015.

[38] "14 advantages of data centralization and standardization in clinical trials." 2017. [Online]. Available: https://goo.gl/p5Nv4J

[39] P. Kumar and H.-J. Lee, "Security issues in healthcare applications using wireless medical sensor networks: A survey," *Sensors*, vol. 12, no. 1, pp. 55–91, 2012.

[40] M. Elhoseny, G. Ramrez-Gonzlez, O. M. Abu-Elnasr, S. A. Shawkat, A. N, and A. Farouk, "Secure medical data transmission model for iot-based healthcare systems," *IEEE Access*, vol. 6, pp. 20 596–20 608, 2018.

[41] B. Yksel, A. Kp, and znur zkasap, "Research issues for privacy and security of electronic health services," *Future Generation Computer Systems*, vol. 68, pp. 1 – 13, 2017.

[42] s. Kirsten, B. Svetlana, B. Rachel Conrad, L. Charles, S. Eliot, and W. Brandon, "Trust and privacy in the context of user-generated health data," *Big Data & Society*, vol. 4, no. 1, 2017.

[43] K. A. Salleh and L. Janczewski, "Technological, organizational and environmental security and privacy issues of big data: A literature review," *Procedia Computer Science*, vol. 100, pp. 19 – 28, 2016.

[44] S. Arora, M. Kumar, P. Johri, and S. Das, "Big heterogeneous data and its security: A survey," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, April 2016, pp. 37–40.

[45] N. Talebi, C. Hallam, and G. Zanella, "The new wave of privacy concerns in the wearable devices era," in *2016 Portland International Conference on Management of Engineering and Technology (PICMET)*, Sept 2016, pp. 3208–3214.

[46] "Library guides: Data resources in the health sciences: Clinical data," 2017. [Online]. Available: http://guides.lib.uw.edu/hsl/data/findclin

[47] P. Jeffery, "Containing the cloud: Security issues in a large scale observational pharmacovigilance research project," *Security and Management*, 2010.

[48] M. Deng, M. Petkovic, M. Nalin, and I. Baroni, "A home healthcare system in the cloud–addressing security and privacy challenges," in *2011 IEEE 4th International Conference on Cloud Computing*, July 2011, pp. 549–556.

[49] M. Meingast, T. Roosta, and S. Sastry, "Security and privacy issues with health care information technology," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2006, pp. 5453–5458.

[50] W. Jie, J. Arshad, and P. Ekin, "Authentication and authorization infrastructure for gridsissues, technologies, trends and experiences," *The Journal of Supercomputing*, vol. 52, pp. 82–96, 2009.

[51] "Ibm: Data driven healthcare organizations use big data analyt-

ics for big." 2018. [Online]. Available: http://www03.ibm.com/
[52] "Top 10 challenges of big data analytics in healthcare." 2017. [Online]. Available: https://goo.gl/ezePbY
[53] S. Rao, S. N. Suma, and M. Sunitha, "Security solutions for big data analytics in healthcare," in *2015 Second International Conference on Advances in Computing and Communication Engineering*, May 2015, pp. 510–514.
[54] P. Jain, M. Gyanchandani, and N. Khare, "Big data privacy: a technological perspective and review," *Journal of Big Data*, vol. 3, no. 1, p. 25, Nov 2016.
[55] Y. Gahi, M. Guennoun, and H. T. Mouftah, "Big data analytics: Security and privacy challenges," in *2016 IEEE Symposium on Computers and Communication (ISCC)*, June 2016, pp. 952–957.
[56] I. Olaronke and O. Oluwaseun, "Big data in healthcare: Prospects, challenges and resolutions," in *2016 Future Technologies Conference (FTC)*, Dec 2016, pp. 1152–1157.
[57] B. Jayant, D. U. Swapnaja, A. A. S. S, and M. Dattatray, "Analysis of dac mac rbac access control based models for security," 2014.
[58] "Sattribute based access control - glossary definition - jericho systems." 2017. [Online]. Available: https://goo.gl/VZkwTJ
[59] M. Evered and S. Bögeholz, "A case study in access control requirements for a health information system," in *Proceedings of the Second Workshop on Australasian Information Security, Data Mining and Web Intelligence, and Software Internationalisation - Volume 32*, ser. ACSW Frontiers '04, 2004, pp. 53–61.
[60] X. Jin, R. Krishnan, and R. Sandhu, "A unified attribute-based access control model covering dac, mac and rbac," in *Proceedings of the 26th Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy*, ser. DBSec'12. Springer-Verlag, 2012, pp. 41–55.
[61] "What are some disadvantages of homomorphic encryption schemes?" 2017. [Online]. Available: https://goo.gl/ZAipPg
[62] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 131–143, Jan 2013.
[63] F. A. Allaert, G. Le Teuff, C. Quantin, and B. Barber, "The legal knowledge of the electronic signature: A key for a secure direct access of patients to their computerized medical record," in *International Journal of Medical Informatics, 73:239242.*
[64] "The nhs plan a plan for investment a plan for reform, presented to parliament by the secretary of state for health," 2017. [Online]. Available: http://www.nhs.uk/nationalplan/nhsplan.htm
[65] R. Zhang and L. Liu, "Security models and requirements for healthcare application clouds," in *2010 IEEE 3rd International Conference on Cloud Computing*, July 2010, pp. 268–275.
[66] F. D. Guilln-GMez, I. Garca-Magario, J. Bravo-Agapito, R. Lacuesta, and J. Lloret, "A proposal to improve the authentication process in m-health environments," *IEEE Access*, vol. 5, pp. 22 530–22 544, 2017.
[67] Y. Zhang, R. H. Deng, G. Han, and D. Zheng, "Secure smart health with privacy-aware aggregate authentication and access control in internet of things," *Journal of Network and Computer Applications*, vol. 123, pp. 89 – 100, 2018.
[68] M. A. Azad, S. Bag, and F. Hao, "Privbox: Verifiable decentralized reputation system for online marketplaces," *Future Generation Computer Systems*, vol. 89, pp. 44 – 57, 2018.
[69] H. Lee, S. Kim, J. W. Kim, and Y. D. Chung, "Utility-preserving anonymization for health data publishing," *BMC Medical Informatics and Decision Making*, vol. 17, no. 1, p. 104, Jul 2017.
[70] L. Zhang, G.-J. Ahn, and B.-T. Chu, "A role-based delegation framework for healthcare information systems," in *Proceedings of the Seventh ACM Symposium on Access Control Models and Technologies*, ser. SACMAT '02. New York, NY, USA: ACM, 2002, pp. 125–134.
[71] D. Gritzalis, "A baseline security policy for distributed healthcare information systems," *Computers & Security*, vol. 16, no. 8, pp. 709 – 719, 1997.