IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 7, 2020, accepted February 19, 2020, date of publication February 27, 2020, date of current version March 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2976687

Key Issues in Healthcare Data Integrity: Analysis and Recommendations

ABHISHEK KUMAR PANDEY[1], ASIF IRSHAD KHAN[2], YOOSEF B. ABUSHARK[2], MD. MOTTAHIR ALAM[3], ALKA AGRAWAL[1], RAJEEV KUMAR[1], AND RAEES AHMAD KHAN[1]

[1]Department of Information Technology, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, Lucknow 226025, India
[2]Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[3]Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Rajeev Kumar (rs0404@gmail.com)

ABSTRACT Managing data integrity is a challenging task for any expert or a researcher. This study attempts to collate a Systematic Literature Review of the research efforts done in the domain of healthcare data integrity. The paper highlights the criticalness of data integrity issues in healthcare through attack statistics in the first section. The second section of the paper systematically reviews the previous studies discussing the healthcare related Systematic literature reviews and data integrity techniques in healthcare sector. The third section of this study examines the collated literature through various analysis methodologies and discusses about the most prioritized technique as well as its challenges in healthcare data security. The fourth section illustrates about the various challenges and future directions to take while constructing a roadmap for the future research endeavors in healthcare data integrity management techniques. The concluding segment of the paper presents an objective assessment and sensitivity analysis for finding the implications and difficulties in the studies while outlining feasible solutions. Furthermore, this research endeavour also conducts a Scientometric analysis of all the studies for better understanding of the literature reviewed. Ranking or the Priority analysis part of the paper is totally dedicated to the previously used techniques in healthcare. The paper also discusses about data integrity techniques and postulates that the most prioritized data integrity technique is the blockchain.

INDEX TERMS Healthcare data, data integrity, data security, systematic review, Scientometric analysis, sensitivity analysis, fuzzy-AHP.

I. INTRODUCTION

Data is the most valuable asset in the current digital era. Every digitalized industry is generating huge amount of data. Managing this big amount of data securely is a complex and challenging task for every security expert. Each type of data has its own significance and use. The importance of data totally depends on its type. For example, healthcare data has its own significance in people's life. As ascertained by the current increase in cyber-attacks, the bad actors are directly targeting this repository of data to exploit the monetary benefits that a pilfered or a tampered data can accrue.

It has been revealed that data integrity rift is often difficult to identify. The importance of data integrity protection has become more significant when the larger ramifications of the data integrity breach are often unknown, as the attackers use the breached data to orchestrate other attacks. In the near future, cyber operations will employ digital information in order to compromise its integrity rather than deleting or disrupting access to it [41]. Even lives will be affected by tampering with information. This emerging kind of cybercrime poses a gigantic threat, which needs to be addressed urgently given the consequences due to data tampering [1]. Therefore, security practitioners and researchers need to be abreast of the perils of data tampering. A continuous and rigorous data protection solution is urgently required in order to guarantee the information protection against manipulation attacks.

Data integrity remains one of the most critical concerns for healthcare industries also. Data integrity breach in healthcare institutions may result in any number of potentially serious consequences. Cybersecurity incidents are now perceived to be the gravest threats to hospitals. Preserving data integrity

The associate editor coordinating the review of this manuscript and approving it for publication was Vlad Diaconita.

in healthcare industries has become a challenging problem because of the organizational structure of the healthcare institutions that entail high-end point complexity and regulatory pressures. Many security breach instances have proved that the healthcare industry is still lagging behind other industries in its efforts to protect the data integrity of its stakeholders.

The integrity of data ensures the organization's brand image and the customers' trust. Any breach in the integrity of data can lead to immense loss of not only the revenue but also dent the customers' trust in the organization's credibility. This kind of threat is more dangerous for organizations as compared to that of attacks on confidentiality and availability. The significance of the data integrity issue becomes even more serious when many attacks of data integrity often remain undetected or unknown and the erroneous information or the data is used by an attacker in different types of attacks.

The *main aim of this Systematic Literature Review (SLR)* is to illustrate the current data integrity techniques that are used by the researchers to secure the healthcare data. The SLR attempted in this study compiles a repository of data integrity techniques that are used by attackers and also tells about those techniques that need in-depth research for better development. This SLR has been envisioned in two phases. In the first phase, the SLR provides brief information about the previous data integrity attacks associated with healthcare industry in order to provide an overview of the criticalness on current data integrity scenarios in healthcare. In the second phase, the SLR provides a systematic review of previous research initiatives related to healthcare data integrity.

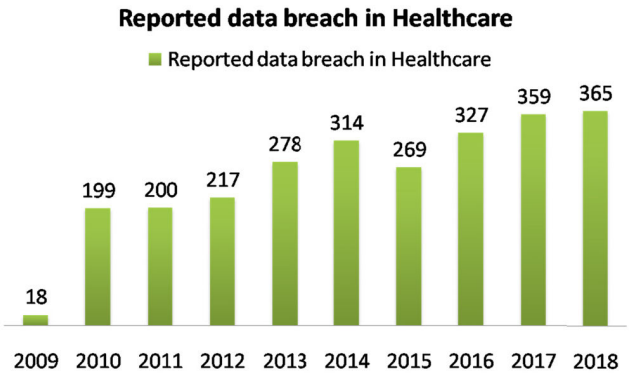

FIGURE 1. Data breaches on the healthcare industry in the last 10 year.

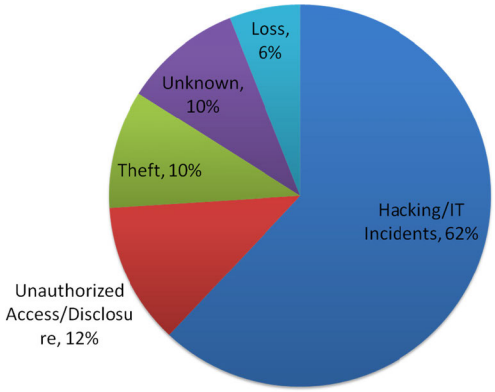

FIGURE 2. Percentage ratio of the type of breach on healthcare industry.

II. CURRENT DATA INTEGRITY RISK PLOT IN HEALTHCAR

Data integrity issue is one of the most demanding concerns for the healthcare industry in the whole world. An integrity breach in a healthcare organization can have disastrous consequences. A patient whose data has been tampered with could be given wrong medications causing fatalities. Most healthcare organizations at present have weak and vulnerable data storage procedures and lack secure mechanisms to foil malware attacks. All these issues create many challenges associated with data integrity in healthcare organizations. Hence, the research team of this study has worked on a novel approach that provides an overview on the current data integrity plot in healthcare through attack statistics. There are numerous surveys which cite the increasing number of data intrusions specifically targeting the healthcare industry. ''HIPPA'', an online survey journal, conducted a study on the data breach attacks on healthcare organizations done during 2009-19. This study shows that in comparison to 2009, the data breach attack on the healthcare industry at present is in its worst condition [2]. The graph of attacks in Figure 1 illustrates that the healthcare industry requires effective safeguards against malware attacks to manage the integrity, confidentiality and availability of data.

HIPPA's report cites 25 largest data breaches in the healthcare industry in the last 10 years. With the help of that record, we have categorized the percentage ratio of the type of attack that was implemented more often in healthcare organizations. Figure 2 shows that 62% of the largest healthcare attacks are implemented by IT incidents alone. This is a big ratio for any industry [2]. Critical analysis of this type of categorization shows the need for a systematic and a foolproof package for managing data integrity and smart hospital security.

According to a study, 94% of the healthcare organizations reported cyber-attacks on their systems [3].An annual analysis report on the breach in healthcare industry tells that the number of breached records tripled in 2018 as compared to 2017 [4]. An online article illustrates that the average cost of any healthcare record on the dark web is from $1 to $1000. This is the second-largest cost for any asset on the dark web [5]. In 2019, 16,819 records of cancer patients were disclosed at Cancer Treatment Centers of America (CTCA), South Eastern Regional Medical Center by targeting their emails [6]. According to an online news website, in early May 2019, American Medical Collection Agency (AMCA) was hacked for 8 months and 25 million patients' data were compromised during this period. Data as classified and sensitive as the billing record, prescription of the patients was compromised during this attack [7].

The recent incidents of data breach reported in two major healthcare industries namely Quest Diagnostics and LabCorp compromised more than 19 million patients' data via a service supplier they shared [8]. According to a new

research report by Global Market Insights, the Global Healthcare Cybersecurity Market is set to surpass USD 27 billion by 2025 [9]. Another shocking case in 2019 is the breach of the data of 10,993 availers in the American Baptist Homes of the Midwest by compromising emails and Network Serves [10].

Statistics discussed in this section of paper clearly explain about the attack trends and provide a review of attacks for healthcare services in previous years. A critical study of these attacks provides a clear status of data integrity and cyber-attacks in healthcare services. Data manipulation also breeds uncertainty. In today's data-driven world, the consequences of uncertainty are frightening. Data integrity breach can undermine the basics of commerce, health, infrastructure national security and political systems. Data manipulation is more insidious, subverting not only the confidence in the ability of an industry to protect its data but also questions the integrity of the industry's data. Imagine the consequences if terrorists manipulate or doctor sensitive military and government data [11]. Manipulation of highly confidential data can lead to catastrophic consequences. This scenario posits the urgent need for understanding the current research status in healthcare data integrity.

Thus, the authors conducted a systematic literature review on previous data integrity techniques employed by researchers and experts. Moreover, the authors have also identified the most effective technique that needs more extensive research and must be a priority of the security experts in their efforts to preserve data integrity.

III. RELATED WORK

In order to conduct a Systematic literature review, the authors surveyed various healthcare related SLR's. Some of these have discussed about the administrative qualities and needs and some are about the various privacy and data security approaches. Authors also found that though healthcare data integrity management is the most crucial and challenging topic for current security experts and researchers, not much literature is available on data integrity issue of healthcare. But whatever survey is available, it gives effective information. It is evident that most researchers have specifically focused on a given data integrity technique or methodology of healthcare in their reviews. The references that this study based its research work on are cited below:

- *P. Asma et al.* provides an exhausting review on big data handling mechanisms. The paper provides a brief and comprehensive knowledge related to big data handling mechanism in healthcare through various aspects. The study categorizes the mechanisms into various fields for an easy and comparative analysis [42].
- *P. P. Biancone et al.* discussed about the healthcare data quality in their paper. Their study illustrates the various data quality assurance methodologies of different research work done and published from 2014 to 2018. The paper chooses various quality research initiatives and analyzes their respective results on various standards [43]. The paper contributes effective information and the current state of data quality methods in healthcare for future researchers.
- *H. M. Hussien et al.* provide a brilliant review on current situation of healthcare for developing a roadmap for blockchain technology. The paper discusses various issues like interoperability, accountability as well as law related implications of healthcare in order to analyze the blockchain technology [44]. The study also describes the roadmap to enable the healthcare industry for blockchain technology and prepare taxonomy. The paper contributes some very significant and effective information for healthcare industry.
- *P. Behrouz et al.* provides a review on Internet of Things (IoT). IoT is the most significant part of smart hospitals. The authors discuss about the data aggregation mechanisms of IoT for better communication and effective use [45]. The study contributes on various aspects as it provides a comprehensive study on various data aggregation mechanisms.

The above discussed research initiatives provide some significant knowledge for healthcare industry through SLR's. However, the authors found that there is a need for a SLR which focuses on various data integrity techniques and provides a roadmap for future researchers to illustrate their research initiatives. In order to achieve this goal, the proposed research endeavor discusses about the various data integrity management techniques that are discussed in top quartile research articles.

IV. LITERATURE EXAMINATIO

Data integrity in healthcare is a sensitive and important issue, yet there is minimal amount of literature review available on this context. Authors have tried to incorporate all quality research initiatives in this SLR and provide a Scientometric analysis of the selected studies. In order to conduct the SLR, the authors have followed the selection methodology that has been discussed in [12]. This SLR is premised on certain key objectives which the authors identified before working on the collation and review of the relevant literature. The section below maps these objectives.

A. RESEARCH OBJECTIVE/PURPOSE

The foremost purpose for selecting the SLR is the need for seeking solutions in the wake of an alarming rise in data breach and manipulation in the healthcare industry. It is necessary and very important to discuss this issue and provide a standalone review for better understanding of researchers and experts. A review always provides an overview of the current situation of the related field. However, in our case, the authors have provided all the required and general information on the field that is needed by a researcher for understanding the issue. The authors have discussed two main and most important objectives of this SLR in this section. These objectives provide a path for the authors to successfully conduct the SLR.

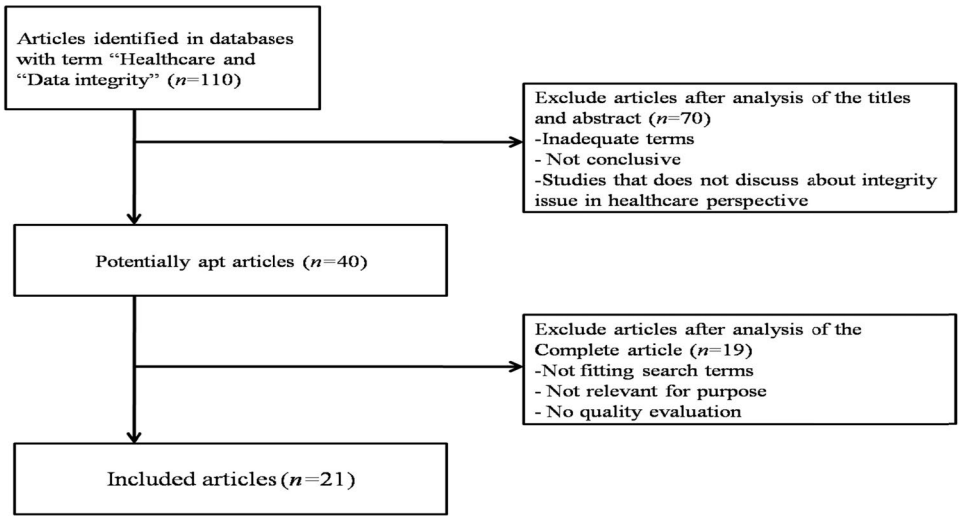

FIGURE 3. Paper selection process.

Following are the Objectives:

Objective 1: What methods were utilized for overseeing information on integrity in past publications for healthcare services?

Motivation: In order to frame workable solutions to stem data breach episodes, it is important to comprehend and then collect the available techniques and methodologies that have already been attempted in this direction. Hence, this SLR attempts to integrate and systematically profile the available literature for an exhaustive reference. Thus, the SLR would be a repository for future researchers to refer to. Furthermore, the biggest motivation for selecting this objective was to draw the attention of the research community towards this immensely critical issue.

Objective 2: Which information integrity method needs utmost concentration in healthcare services?

Motivation: Authors of this study intended to provide a prioritization list of data integrity techniques according to their need of focus which would help the future researchers most. Prioritizing the previous studies would also help the future researchers in selecting the most effective approach and in understanding the need for the healthcare sector.

B. METHODOLOGY

In order to conduct the SLR, authors tried to include only data integrity-related papers. For this, the scientific data repository PubMed, Science Direct, IEEE Xplorar and Google Scholar were used to collect the relevant studies. Furthermore, following keywords are used for the search *Healthcare, Data Integrity, Data Security, Secure Data Sharing* with Boolean operator AND. 110 studies were identified at the initial level and after applying different exclusion filters, the authors found 21 relevant studies for conducting the healthcare data integrity SLR. To find good and effective literature conclusions, the authors set some inclusion and exclusion criteria, which are discussed below.

For including the papers authors followed the following criteria:

- Authors have included the studies that discuss about data integrity in healthcare as a security problem and provide some empirical solution.
- SLR includes papers that use a specific approach for resolving the integrity issue in healthcare.
- SLR only includes studies that are published in Q1 and Q2 journals (for result accuracy and validation).
- SLR includes studies that provide some conclusive result on integrity issue of healthcare.

For excluding the papers, the authors applied the following criteria:

- Exclude papers that were not relevant to the search terms and purpose of review.
- Exclude papers that discussed data integrity issue but not from healthcare perspective.
- Exclude papers that were are not effective and conclusive in order to help the data integrity issue in healthcare.

As described in figure 3, in the first phase, the authors excluded the papers on the basis of their titles and abstracts.70 studies that were not relevant to this SLR were excluded in the first phase itself.

In the second phase, the authors excluded the papers after analyzing the complete article. In this phase 19 studies that were not suitable for SLR were excluded. Authors use Preferred Reporting Items for Systematic Literature Review and Meta-Analysis (PRISMA) 2009 Flow Diagram for illustrating the paper selection process. This method was introduced by [13] and provides rules to create Systematic Reviews and Meta-examination. Table 1 describes the tabulated form of studies and their respective relevant percentage.

Table 1 describes the different search figures from various digital libraries for better understanding of the search procedure. The authors' aim is to describe the current data integrity techniques that are used by healthcare organizations

TABLE 1. Search figures.

Data Repository	Relevant Papers	Not-Relevant Papers	Total	Relevant Percentage (%)
PubMed	5	40	45	11.11
Science Direct	3	99	102	2.94
Google Scholor	2	598	600	0.33
IEEE Xplorar	8	3437	3445	0.23
Cite-Seeker	3	17	20	15.00
Total	21 (After removing 19 duplicates from Google scholar)	4191	4212	

through a systematic review and highlight the criticalness of data integrity issues in healthcare. This has been done by collating data on various breach statistics.

C. EXPLORATORY ANALYSIS OF RESULTS

An essence of the included studies is shown in table 2. The table shows the main content of the studies and their corresponding data integrity approach. During the review it was found that some papers also discussed the integrity and healthcare challenges in their content. Hence for a more comprehensive review, the authors included these papers for this SLR. The description of different data integrity techniques that were used in previous studies is tabulated below.

1) BLOCKCHAIN APPROAC

Many researchers have used the blockchain approach as a key attribute in their studies for managing healthcare data securely and these include:

- A paper proposed by William J. Gordon et al. discusses how we facilitate the blockchain approach in a healthcare organization [14]. The paper discusses the challenges and issues and proposes a novel model for facilitating the blockchain methodology in healthcare services.
- Peng Zhang et al. presented a paper that discussed the clinical data security and provided blockchain-based architecture FHIR-Chain [16].
- Abdullah Al Omar et al. presented a paper discussing blockchain as storage in healthcare [26].
- Anastasia Theodouli et al. presented a novel blockchain approach for facilitating healthcare data auditable, sharable and securely usable [28].
- Xueping Liang et al. proposed a model for personalized healthcare data for secure sharing and a decentralized blockchain approach for enhancing the security of blockchain in healthcare services [29].

2) MASKED AUTHENTICATED MESSAGING EXTENSION

James Brogan et al. presented a paper discussing the security improvement of healthcare data through masked authentication messaging extension module in wearable medical devices [15]. The authors established a relationship between IOAT and masked authentication messaging extension in the paper and solved the challenges that are associated with wearable devices. The approach that is used in this paper is very useful for future researchers.

- *Secure-BSN:* Prosanta Gope et al. presented a paper discussing the Body sensor network approach in healthcare IoT environment. BSN approach is core technology in healthcare IoT environment where a patient is monitored through tiny light-weight body sensors [18]. The paper provides a secure and integrity manageable BSN approach for secure IoT communication in healthcare organizations.
- *Authentication:* P. Vimala Chandran et al. proposed an authentication step in Australian healthcare services [19]. The authentication step provides a significance tracking of data for patients and they can control the access of data through this novel approach.
- *Encryption:* M. ELHOSENY et al. presented a paper discussing the security of healthcare images and patients reports in image format and other types. The paper provides steganography and a hybrid encryption mechanism for securing healthcare data. The approach needs more research for better results in future [20].
- *Wolf-Coding-Based Secret Sharing:* EntaoLuo et al. provide a detailed secure IoT communication and data sharing between two IoT devices [21].The authors used a Wolf-coding-based sharing methodology for securing the IoT environment in healthcare.
- *Secure Cloud:* Gunasekaran Manogaran et al. presented a paper discussing the big data scenario in current healthcare sector and provide a secure cloud approach for managing big healthcare data [23]. Benjamin Fabiana et al. also discuss the inter-organizational data transfer through secure cloud approach.
- *Merkle Tree-Based Approach:* Brihat Sharma et al. presented a paper providing an approach that is used for secure data transfer and communication in healthcare services [32]. The proposed approach mimics blockchain approach and tries to provide a better and secure environment for data transfer and communication.

D. UNIT ANALYSIS

The unit analysis is part of systematic literature review in which the authors describe and categorize the studies according to their corresponding subfield of healthcare. For example if a study is providing full integrity managed system for whole healthcare system then the sub-category that is defined in this SLR is ''Whole Healthcare System'' and if a study only covers the secure communication between IoT devices then the sub-category of that field is Data transfer. Table 3 illustrates the different studies covering different aspects of the healthcare system for managing data integrity.

Table 3 describes the current research studies that have focused on different aspects of healthcare sector. The table clearly shows that enhancing medical data integrity requires more significance in comparison to the other aspects of healthcare services. A strong integrity managed mechanism

TABLE 2. Exploratory analysis of studies.

Author	Study Description	Technique for Data Integrity
William J. Gordon et al. (2018) [14]	The study provides descriptive information of how to facilitate the block chain approach in healthcare sector. The paper also discusses about the challenges that are associated with blockchain in order to provide a secure communication.	Blockchain
James Brogan et al. (2018) [15]	The study provides distributed Ledger technologies in advancing electronic health information's. The paper provides a cost-effective and novel approach for the healthcare organization.	Masked Authenticated Messaging extension
Peng Zhang et al. (2018) [16]	The paper provides a blockchain-based architecture FHIR-Chain for securing Medicare.	Blockchain
Christian Esposito et al. (2016) [17]	The study uses cloud storage environment for data available in healthcare organizations and for patients. Authors also use blockchain approach for secure lab report transaction and communication.	Blockchain
ProsantaGope et al. (2015) [18]	The study uses Body Sensor Network approach for facilitating secure and integrity managed architecture of IoT in healthcare.	Secure-BSN
P. Vimalachandran et al. (2017) [19]	Authors proposed authorization based model for Australian healthcare services.	Authentication
M. ELHOSENY et al. (2018) [20]	The study provides a stenographic technique with hybrid encryption mechanism for securing health records and images.	Encryption
EntaoLuo et al. (2018) [21]	The study provides a secure sharing based data transfer in IoT environment for data security of healthcare organization.	Slepian-Wolf-coding-based secret sharing (SW-SSS)
Moshaddique Al Ameen et al. (2010) [22]	The paper discusses about the challenges and issues associated with the wireless sensors in healthcare sector.	-
Gunasekaran Manogaran et al. (2017) [23]	Authors give a secure Organizational IoT based model for storing and processing wearable sensor data in medical services.	Secure Cloud
Benjamin Fabiana et al. (2014) [24]	The study provides inter organizational data transfer security through various security attributes. The paper provides the architecture for secure data transfer from one organization to another.	Secure Cloud
Jinyuan Sun et al. (2011) [25]	The paper provides a secure health record system for Patient Privacy based on cryptographic techniques and IoT environment of healthcare industry.	Cryptography
Abdullah Al Omar et al. (2017) [26]	The study presents a data management system for healthcare services to facilitate patients through blockchain technology.	Blockchain
Sue Bowmanet et al. (2013) [27]	The study highlights the current challenges and other error causes in healthcare data integrity in healthcare organization. The paper provides a review on current HER system of healthcare.	-
Anastasia Theodouli et al. (2018) [28]	The study presents mechanism for facilitating blockchain technology for providing auditable and sharable data in healthcare organization.	Blockchain
Xueping Liang et al. (2017) [29]	The study suggests a unique medical data mutual sharing platform by providing an access controlled blockchain to protect private data and enhance the user management using the blockchain.	Blockchain
Karim Abouelmehdi et al. (2018) [30]	In this study, the authors have discussed about the challenges and survey the current situation of healthcare big data.	-
AnamSajid et al. (2016) [31]	The study presents review on healthcare medical data security for providing privacy to the patients. Paper also discusses about the currently used techniques and approaches in healthcare system.	-
Brihat Sharma et al. (2018) [32]	The study proposes a model, the Merkle tree-based approach to secure the integrity of health records. The software model closely refers to the Blockchain technology.	Merkle tree-based approach
Katharine Gammon (2018) [33]	The article illustrates the blockchain application in healthcare sector in various domains.	-

for whole healthcare system is also required through various data integrity management techniques.

E. SCIENTOMETRIC ANALYSIS

In the third step for understanding which data integrity technique must be given more research interest, the authors performed the Scientometric analysis. A scientometric analysis is a quantitative and qualitative analysis of studies. This concept was firstly created by [34].The results of the studies on both quantitative and qualitative analysis have been summarized in table 4 along with the authors, Journal indexed in, ranking, category and quartile. The quartile field reflects all the categories that the journals have according to their Indexed classification.

Table 4 clearly illustrates that maximum number of publications are in Computer Science category. A total of 5 publications are available in Computer science category. Informatics and health information categories have 2-2 publications, respectively. Medicine (miscellaneous) has 2 publications in its category. The engineering category also has 2 publications. All these statistics show that interest of researchers is growing comparatively high in computer science for solving the problem of data integrity in healthcare sector.

All journals have published only one paper except the Journal of Computational and Structural Biotechnology. CSB journal published 3 papers on data integrity. Quartiles of paper strictly show that research work quality

TABLE 3. Unit analysis.

Authors	Healthcare System	Data Transfer	Data Sharing	Data Storage	Patient Data
William J. Gordon et al. (2018)		✓	✓		
James Brogan et al. (2018)		✓			
Peng Zhang et al. (2018)			✓		
Christian Esposito et al. (2016)		✓	✓	✓	
ProsantaGope et al. (2015)	✓				
P. Vimalachandran et al.	✓				
M. ELHOSENY et al. (2018)					✓
EntaoLuo et al. (2018)			✓		
Moshaddique Al Ameen et al. (2010)			✓		✓
GunasekaranManogaran et al.			✓	✓	
Benjamin Fabiana et al. (2014)		✓			
Jinyuan Sun et al. (2011)					✓
Abdullah Al Omar et al. (2017)				✓	
Sue Bowman et al. (2013)		✓		✓	
Anastasia Theodouli et al.		✓	✓		
Xueping Liang et al. (2017)		✓		✓	
Karim Abouelmehdi et al. (2018)	✓		✓		
AnamSajid et al. (2016)		✓		✓	
Brihat Sharma et al. (2018)					✓
Katharine Gammon (2018)					

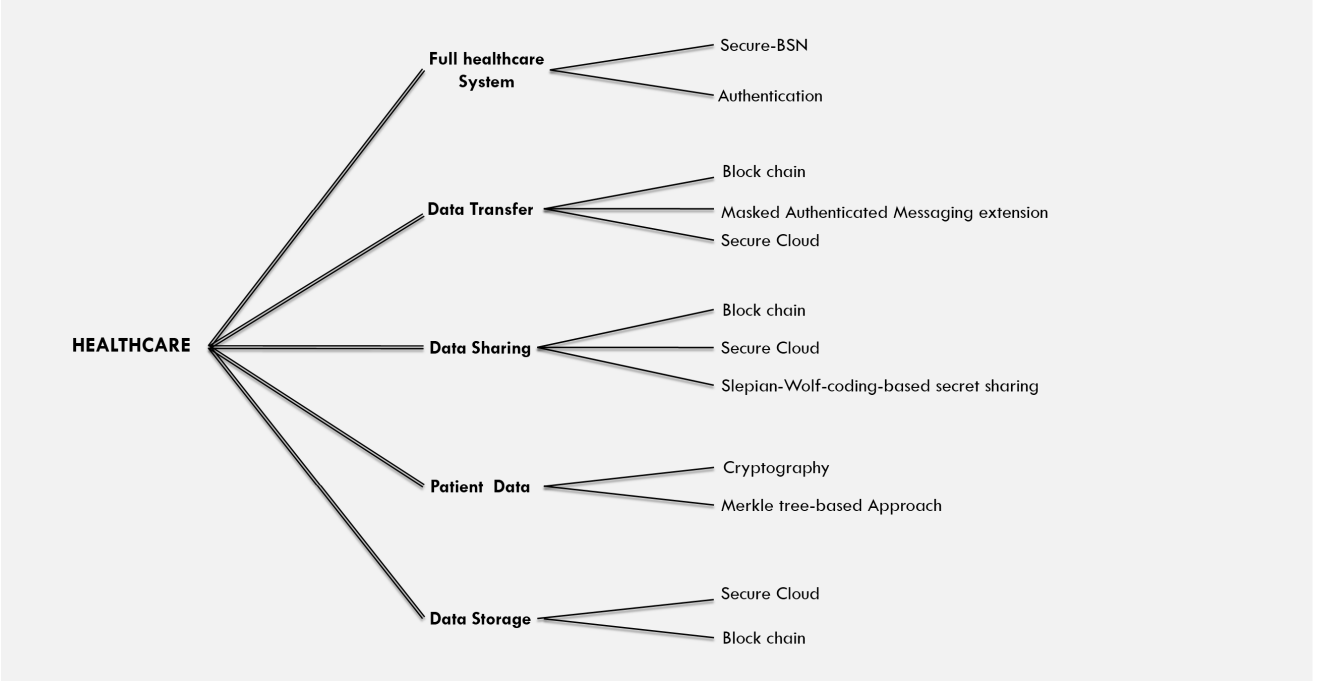

FIGURE 4. Hierarchy of data integrity techniques used in different sub-fields of healthcare.

TABLE 4. Scientometric analysis.

Authors	Journal	Quartile	Category
William J. Gordon et al. (2018)	Computational and Structural Biotechnology Journal	Q1	Computer Science Application
James Brogan et al. (2018)	Computational and Structural Biotechnology Journal	Q1	Computer Science Application
Peng Zhang et al. (2018)	Computational and Structural Biotechnology Journal	Q1	Computer Science Application
Christian Esposito et al. (2016)	IEEE Cloud Computing	Q1	Computer Science (miscellaneous)
ProsantaGope et al. (2015)	IEEE SENSORS JOURNAL	Q1	Electrical and Electronical Engineering
P. Vimalachandran et al.	2017 International Conférence on Orange Technologies (ICOT)	-	-
M. ELHOSENY et al. (2018)	IEEE Access	Q1	Engineering (miscellaneous)
EntaoLuo et al. (2018)	IEEE Communications Magazine	Q1	Computer Networks and Communications
Moshaddique Al Ameen et al. (2010)	Journal of Medical Systems	Q2	Health Informatics
GunasekaranManogaran et al.	Thames L., Schaefer D. (eds) Cybersecurity for Industry 4.0. Springer Series in Advanced Manufacturing	-	-
Benjamin Fabiana et al. (2014)	Information Systems	Q1	Information System
Jinyuan Sun et al. (2011)	2011 31st International Conference on Distributed Computing Systems	-	-
Abdullah Al Omar et al. (2017)	International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage	-	-
Sue Bowmanet al. (2013)	Perspective Health Information Managing	Q2	Medicine (miscellaneous)
Anastasia Theodouli et al.	2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering(TrustCom/BigDataSE)	-	-
Xueping Liang et al. (2017)	2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)	-	-
Karim Abouelmehdi et al. (2018)	Journal of Big Data	Q1	Information System and Management
AnamSajid et al. (2016)	Journal of Medical Systems	Q2	Health Information Management
Brihat Sharma et al. (2018)	2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)	-	-
Katharine Gammon (2018)	Nature Medicine	Q1	Medicine (miscellaneous)

is very high in healthcare data integrity techniques but there is shortage of research work in this field. For achieving the goal of manipulation free data management in healthcare services, it is strictly recommended to develop and conduct high-quality research on continuous basis.

TABLE 5. Fuzzy pair-wise comparison matrix.

	(C1)	(C2)	(C3)	(C4)	(C5)
(C1)	1.0000,1.0000,1.0000	1.0000, 1.5000, 1.9000	0.5000, 0.6000, 1.0000	0.4000, 0.6000, 1.0000	0.2000, 0.3000, 0.4000
(C2)	-	1.0000, 1.0000, 1.0000	0.6000, 0.7000, 0.8000	0.3000, 0.4000, 0.6000	0.3000, 0.4000, 0.5000
(C3)	-	-	1.0000, 1.0000, 1.0000	1.0000, 1.3000, 1.6000	0.3000, 0.4000, 0.8000
(C4)	-	-	-	1.0000, 1.0000, 1.0000	0.5000, 0.9000, 1.6000
(C5)	-	-	-	-	1.0000, 1.0000, 1.0000

TABLE 6. Fuzzy aggregated pair-wise comparison matrix at level 2 for full healthcare system.

	(C11)	(C12)
(C11)	1.0000, 1.0000, 1.0000	0.7000, 0.9000, 1.1000
(C12)	-	1.0000, 1.0000, 1.0000

TABLE 7. Fuzzy aggregated pair-wise comparison matrix at level 2 for data transfer.

	(C21)	(C22)	(C23)
(C21)	1.0000, 1.0000, 1.0000	0.7000, 1.2000, 1.7000	0.7000, 1.0000, 1.4000
(C22)	-	1.0000, 1.0000, 1.0000	1.2000, 1.6000, 2.2000
(C23)	-	-	1.0000, 1.0000, 1.0000

F. RANKING/PRIORITY ANALYSIS

The above description of studies categorizes the previous studies of data integrity techniques in healthcare into different standards for easy and clear understanding of the previous scenario. The authors have added a ranking analysis methodology using an effective Fuzzy-Analytical Hierarchy Process (AHP) for prioritizing the data integrity techniques and provided the highest ranked technique for the research community.

For analyzing the previous studies and applying AHP, authors have a created a hierarchy of data integrity techniques covering different sub-fields of healthcare system. Figure 4 shows the hierarchy of integrity techniques in different healthcare domains.

The above hierarchy describes various data integrity techniques that are used in different sub-fields of healthcare system. Authors have applied the Fuzzy-AHP methodology for assessing the priority of the data integrity techniques.

1) PRIORITY ASSESSMEN

Fuzzy-AHP technique is good in providing crisp and accurate decisions [35]. Fuzzy-AHP is a widely used priority assessment tool. Authors have also used this tool to assess the most prioritized data integrity methodology for healthcare. This type of classification and decision provides a novel and valuable idea to the future researchers.

In order to conduct the Fuzzy-AHP methodology, the authors have used this technique previously [36] and collected data from 75 experts from different fields. With the help of inputs from the experts, this methodology aims to provide the most prioritized data integrity technique in healthcare. Figure 4 shows the hierarchy of the data integrity techniques that are used in healthcare. With the help of [35], [36], the constructed and aggregated fuzzy comparison metrics have been prepared.

Table 5 shows the fuzzy pair-wise comparison matrix of level one. Level 1 includes Full Healthcare System, Data Transfer, Data Sharing, Patient Data and Data Storage. Table 6 shows fuzzy pair-wise comparison matrix at level 2 of full healthcare system. Level 2 of full healthcare system contains Secure BSN and Authentication data integrity techniques. Table 7 shows the fuzzy pair-wise comparison matrix at level 2 of Data transfer. Level 2 of data transfer contains Blockchain, Masked Authenticated message extension and Secure Cloud. Table 8 shows the fuzzy pair-wise comparison matrix at level 2 of data sharing. Level 2 of data sharing includes Blockchain, Secure Cloud and Slepian-wolf coding based secret sharing. Table 9 shows the fuzzy pair-wise comparison matrix of level 2 for patient's data domain. Level 2 for patient's data contains Cryptography and Markel tree based approach. Table 10 shows the fuzzy pair-wise comparison matrix of level 2 for Data Storage. Level for Data

TABLE 8. Aggregated pair-wise comparison matrix at level 2 for data sharing.

	(C31)	(C32)	(C33)
(C31)	1.0000, 1.0000, 1.0000	0.2000, 0.3000, 0.4000	.2000, 1.4000, 1.7000
(C32)	-	1.0000, 1.0000, 1.0000	.3000, 0.4000, 0.6000
(C33)	-	-	.0000, 1.0000, 1.0000

TABLE 9. Aggregated pair-wise comparison matrix at level 2 for patient data.

	(C41)	(C42)
(C41)	1.0000, 1.0000, 1.0000	1.2000, 1.4000, 1.7000
(C42)	-	1.0000, 1.0000, 1.0000

TABLE 10. Aggregated pair-wise comparison matrix at level 2 for data storage.

	(C51)	(C52)
(C51)	1.0000, 1.0000, 1.0000	0.3000, 0.4000, 0.6000
(C52)	-	1.0000, 1.0000, 1.0000

TABLE 11. Defuzzified pair-wise comparison matrix and local weights.

	(C1)	(C2)	(C3)	(C4)	(C5)	Weights
(C1)	1.0000	1.5000	0.7000	0.6000	0.3000	0.1315
(C2)	0.7000	1.0000	0.7000	0.4000	0.4000	0.1094
(C3)	1.4000	1.5000	1.0000	1.3000	0.5000	0.1975
(C4)	1.6000	2.4000	0.8000	1.0000	1.0000	0.2356
(C5)	3.3000	2.7000	2.0000	1.0000	1.0000	0.3260
						C.R.= 0.00245

TABLE 12. Aggregated pair-wise comparison matrix at level 2 for full healthcare system.

	(C11)	(C12)	Weights
(C11)	1.0000	0.9000	0.4737
(C12)	1.1000	1.0000	0.5263
			C.R.=0.0002

storage contains Secure Cloud and Blockchain methodology. To evaluate the defuzzify values and CR values of the matrix, authors adopted [36] processes. For defuzzification process, this paper used α cut method [36]. Table 11 to Table 16 shows the defuzzified pair-wise comparison matrix. Local weights of each group are also shown in the Table 11 to Table 16. Finally, dependent weights through the hierarchy have been shown in Table 17.

Table 17 discusses about the results determined after the calculation of data integrity methods through fuzzy-AHP technique. The table shows that blockchain technique has the highest priority ranking amongst all the techniques. The findings corroborate that the researchers need to focus on the blockchain methodology for better integrity management approaches and environment according to the fuzzy-AHP model. For more information and description, the authors have discussed about the blockchain challenges in healthcare. Further, previous blockchain studies done in the context of these challenges have also been discussed. This kind of classification would provide crystal clear information on the current scenario of blockchain research done for data integrity in the healthcare industry.

G. COMPARATIVE STUDY OF PREVIOUS BLOCKCHAIN STUDIES IN RESPECT TO BLOCKCHAIN CHALLENGES IN HEALTHCARE

A descriptive review [35] on blockchain technology is available for understanding the current scenario of blockchain. The paper [35] describes the whole blockchain technology and reviews the associated challenges therein. Authors of this SLR also resourced some healthcare related blockchain

TABLE 13. Aggregated pair-wise comparison matrix at level 2 for data transfer.

	(C21)	(C22)	(C23)	Weights
(C21)	1.0000	1.2000	1. 0000	0.3521
(C22)	0.9000	1.0000	1.6000	0.3647
(C23)	1.0000	0.6000	1.0000	0.2832
				CR=0.0230

TABLE 14. Aggregated pair-wise comparison matrix at level 2 for data sharing.

	(C31)	(C32)	(C33)	Weights
(C31)	1.0000	0.3000	1.4000	0.2445
(C32)	3.5000	1.0000	0.4000	0.3594
(C33)	0.7000	2.4000	1.0000	0.3961
				C.R.= 0.00354

TABLE 15. Aggregated pair-wise comparison matrix at level 2 for patient data.

	(C41)	(C42)	Weights
(C41)	1.0000	1.4000	0.5833
(C42)	0.7000	1.0000	0.4167
			C.R.= 0.0005

TABLE 16. Aggregated pair-wise comparison matrix at level 2 for data storage.

	(C51)	(C52)	Weights
(C51)	1.0000	0.4000	0.2857
(C52)	2.4000	1.0000	0.7143
			C.R.=0.004

TABLE 17. Overall weights and ranking of methods.

First Level Methods	Local Weights of First Level	Second Level Methods	Local Weights of Second Level	Overall Weights	Percentage	Overall Ranks
C1	0.1315	C11	0.4737	0.0623	6.23 %	8
		C12	0.5263	0.0692	6.92 %	7
C2	0.1094	C21	0.3521	0.0385	3.85 %	11
		C22	0.3647	0.0399	3.99 %	10
		C23	0.2832	0.0310	3.10 %	12
C3	0.1975	C31	0.2445	0.0483	4.83 %	9
		C32	0.3594	0.0710	7.10 %	6
		C33	0.3961	0.0783	7.83 %	5
C4	0.2356	C41	0.5833	0.1374	13.74 %	2
		C42	0.4167	0.0982	9.82 %	3
C5	0.3260	C51	0.2857	0.0931	9.31 %	4
		C52	0.7143	0.2329	23.29 %	1

challenges and tried to do a comparative analysis between the present challenges and the previous studies. The two main challenges of blockchain that are associated with healthcare domain are:-

1) SCALABILITY RELATED CHALLENG

The primary use of blockchain technology can be related to financial transactions. The above studies show that the tremendous use of blockchain in healthcare field started after

2015. Basic architecture of blockchain is developed to carry small amount and size of data over a block. But the amount of data that is used in a healthcare sector is vast and larger than the amount of data carried by a blockchain.

2) PRIVACY-LEAKAGE RELATED CHALLENG

Fundamental functionality of blockchain approach works on a distributed network environment. Every participant in a network has a replica or copy of transaction. This type of scenario creates many privacy leakage related challenges for blockchain approach. Privacy is a core challenge that is associated with blockchain technology. Many researchers are focusing on the privacy concern of the blocks.

Table 18, below, shows the studies that have discussed about the blockchain challenges and the table also illustrates the implementation status of the previous studies.

Table 18 describes that 80% of the studies do address the healthcare related challenges of blockchain. Implementation status of studies shows that there is a need for practical work in the field of blockchain. The above table also provides a current scenario of blockchain in healthcare data integrity for future researchers.

Priority assessment and comparative study of previous blockchain healthcare publications provides a path for researchers and experts to identify the actual status of blockchain technique as a healthcare data integrity methodology.

V. CHALLENGES & FUTURE DIRECTIONS

After analyzing the studies selected for this SLR, the authors found that there are several issues associated with the effective deployment of Data Integrity management techniques in healthcare. These challenges are mapping future directions and unbolting new doors for researchers to conduct their research and provide inventive but viable solutions for data integrity management systems especially pertaining to healthcare.

For better understanding and easiness, authors have categorized the challenges into different categories.

A. CHALLENGES ASSOCIATED WITH FUTURE OF DATA INTEGRITY BREACHES IN HEALTHCARE

Breach of data integrity in healthcare is a new trend for attackers. The high commercial value of healthcare information makes it an attractive revenue generation sector for attackers. Previous attack statistics discussed in the first section of this paper underline the criticalness of data integrity and data breach in healthcare sector. For understanding the future of attacks and future criticalness of data integrity issue in healthcare, authors conducted a forecasting of data breaches in healthcare based on previous available data. Figure 5 shows the forecasted scenario of data breaches in healthcare.

Figure 5, above, depicts that the data breach scenario in next 10 years is going to be worst in comparison of 2009. This kind of mercurial rise in data breach attacks creates many challenges for future researchers who should work on possible solutions. Software engineering field of computer science works on "*Early Detect, Early Solve*" and for better results in data integrity in healthcare, authors of this study also strictly recommend the same. Prediction of attacks will help the researchers to find the complexity and necessity of this topic in research and, further, motivate them in finding better solutions for the data integrity issue in healthcare. Forecasting of healthcare data breach attacks would enable the experts and the researchers in ensuring prompt and better procurement as well as in establishing prevention mechanisms.

B. CHALLENGES ASSOCIATED WITH BLOCKCHAIN APPROACH

Blockchain approach is typically used for financial transactions from the beginning. The use of blockchain in healthcare data handling is a critical and complex job for any researcher. Many researchers [14], [16], and [17] tried to solve the issues and challenges, yet lot remains to be streamlined. The biggest challenge that is associated with blockchain approach is its data storage capacity in particular block, i.e., the volume issue. Healthcare data is a type of big data and the capacity of blockchain technology is familiar with financial data that is very low in volume. This type of challenge can cause data collapse in between the transaction and mishandling of data. Besides this, blockchain has privacy-related issues as well. In a blockchain environment, data is stored on a distributed network and every node of that network has a copy of transaction for validation purpose. This type of complex structure creates a manipulation free environment but it also creates a challenge for experts to manage the privacy of data in blockchain environment.

C. CHALLENGES ASSOCIATED WITH HEALTHCARE DATA RISK

From the above analyses and their relevant results it is fully evident that many researchers and experts are working on healthcare data integrity as well as data security. While various data security techniques and approaches are available and have been implemented in healthcare organizations, it is also clearly evident that there has been no decline in the proportion of healthcare attacks (from second section of this paper). In addition, the security process of health data in web applications involves many aspects. It includes a database security domain as well as application-layer security and client-based security domains. The authors of this study have analyzed some effective and important issues that directly harm the security of health data in web applications. These issues are written following:

- *Frequent DDoS Attacks:* In the process of analysis, the authors found that there has been a large and challenging growth in DDoS attacks in comparison to previous years. A report stated that in 2018, 1TBPS sized DDoS attack was reported [37]. DDoS attacks have their significant impact on healthcare data. Assume a database server is under a DDoS attack, it can put any

TABLE 18. Previous studies on blockchain challenges.

Author	Scalability challenge	Privacy related challenge	Implementationof work
William J. Gordon et al. (2018) [14]	✓	✓	NO
Peng Zhang et al. (2018) [16]	✓	✓	YES
Christian Esposito et al. (2016) [17]		✓	NO
Abdullah Al Omar et al. (2017) [26]		✓	YES
Anastasia Theodouli et al. (2018) [28]	✓	✓	NO
Xueping Liang et al. (2017) [29]	✓	✓	NO

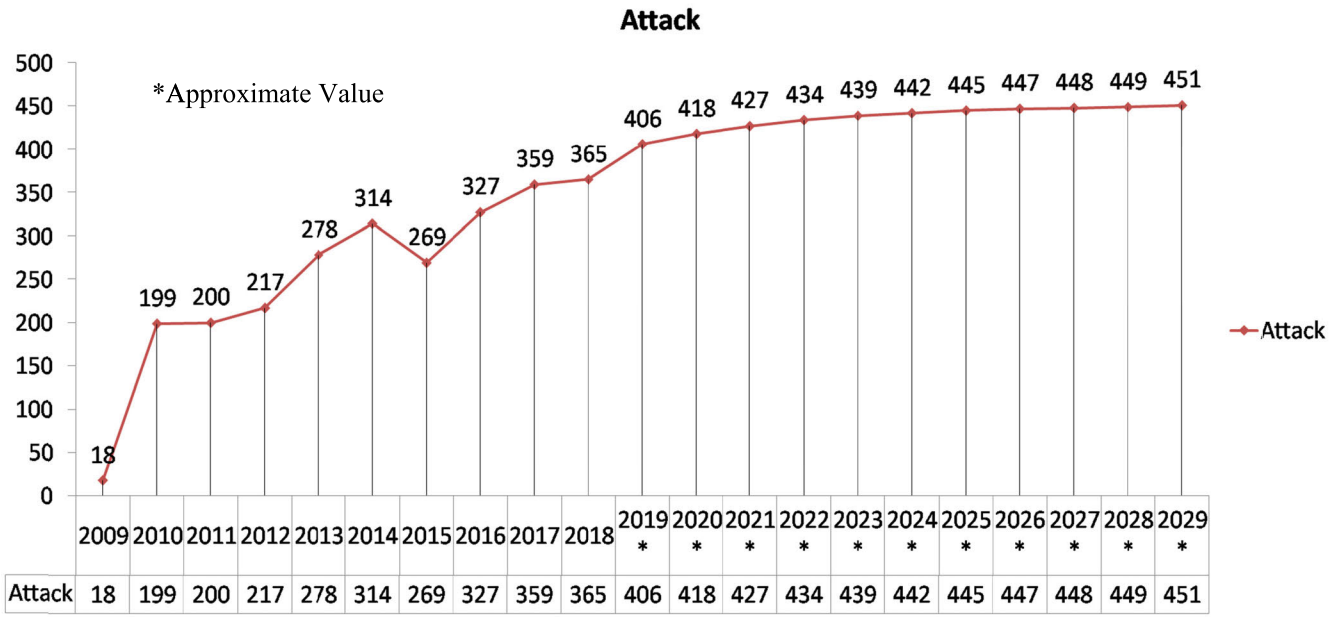

FIGURE 5. Forecasting of data breach.

healthcare organization as well as patients in a life-threatening state. No emergency or daily service is available for that period and this can cause many serious revenue-related as well as life-related losses.

- *Advance Persistent Threat:* APT (Advance Persistent Threat) attacks are the new and most effective technique of intrusions in the modern era. APT is a type of attack title or term that is used when an attack is executed via an intruder or through a group of intruders for a long time in any type of system and network [38]. These types of attacks are most dangerous and harmful for healthcare industry. APT attacks can prey on the healthcare networks and web application securities like a termite and render them ineffective. Authors strictly suggest that there is a need for advanced security measures for preventing APT attacks in healthcare sector.
- *SQL Security Vulnerabilities:* Database security is the most prioritized and significant field for any healthcare organization. It is clearly evident from previous attack statistics that intruders target databases directly for instant mining of information. In between this type of situation a hole in SQL can cause many issues in healthcare data security. A website called *CVE Details* has released an exhaustive list of SQL vulnerabilities that can cause serious harm to any healthcare web application [39].
- *Client Side Vulnerabilities:* This type of issue has various factors that affect the exploitation like the human error about which we have talked in the next heading, phishing or spoofing contents, etc. Spoofing or phishing is a tricky attack that is executed through social engineering on a target. Spoofing attack tricks the victim to download or visit some malicious application or link that can cause exploitation in the system [40]. It is mostly seen in healthcare attacks where the intruders use some phishing websites and contents to fool the employee or victim and exploit the system. All these issues discussed above are affecting the healthcare data security.

D. CHALLENGES ASSOCIATED WITH HUMAN AWARENESS

Many papers have dwelt on the human errors [19] and lack of required level of awareness in healthcare sector. Human awareness in healthcare organizations is the primary need because it is the users who will be facilitated by secure technologies and approaches in a healthcare organization. For managing human errors and awareness there is a need for strong data integrity policies and rules for healthcare system. Policy development for managing data integrity in healthcare organizations is a key topic for future researchers. A Strong policy provides an organization with a better and a healthy environment.

E. CHALLENGES ASSOCIATED WITH IMPLEMENTING NEW TECHNOLOGIES IN THE HEALTHCARE SECTO

Table 4 in this SLR denotes that there is quality research available in data integrity techniques for healthcare sector. But the ground reality of healthcare sector shows that there are still gaps in managing updated data integrity techniques in real environment. It is very important for the organizations to adopt new approaches for better security and management of data in healthcare.

F. CHALLENGES ASSOCIATED WITH RESEARCHERS

The above discussion and findings emphasize that there is a need for more quality and empirical research to ensure that systems become more secure. As discussed by the authors in the second section of this proposed paper, many SLRs are discussing specific data integrity techniques and factors [46] that affect modern health data security. But healthcare lacks the SLR discussing the whole aspect of data integrity techniques and provides an excellent review of it. This SLR is an exhaustive reckoner on the new approaches cited by some researchers which can be adapted as the basis for further research endeavors. There are many papers/articles on the relevant topic in literature. But, for impactful findings, this SLR focuses on only Q1 and Q2 Quartiles.

VI. DISCUSSION

This section underlines the assessment of the objectives that had motivated the authors to profile the SLR for addressing the concern of data breaches in healthcare.

A. ASSESSMENT OF OBJECTIVE 1

Authors chose the descriptive analysis of previous studies as their first objective for this SLR. Authors categorized and analyzed the previous studies on different analysis standards as exploratory analysis, unit analysis and scientometric analysis. These analysis categories would help the readers to understand the previous research scenario of data integrity in healthcare. Exploratory analysis provides details of previous studies like which study is based on what type of data integrity management and other information. Unit analysis may help the researchers to understand that which field of healthcare is covered under which study for assessing the previous research interest. Lastly, the scientometric analysis may provide information related to previous studies quality and publication information. This type of information helps the researchers in preparing their research work and motivating them for good quality work. All these three analysis categories helped the authors in achieving the first objective of the study.

B. ASSESSMENT OF OBJECTIVE 2

Prioritizing the previous data integrity techniques of healthcare is the second objective of authors for this SLR. In order to achieve this objective, the authors conducted a ranking analysis section in this paper. As per the knowledge of the authors no other healthcare data integrity related paper has conducted a prioritizing analysis for data integrity techniques. The Fuzzy-AHP ranking methodology has been enlisted for assessing the rank of data integrity techniques in healthcare. Authors prepared a hierarchy of data integrity techniques in healthcare and applied the fuzzy-AHP methodology on this hierarchy for assessment. This type of analysis in a systematic review provides a novel and clear direction to the future researchers in data integrity techniques. Ranking assessment recommends that researchers must focus on the blockchain methodology.

Hence, both the stated objectives have been realized. Furthermore, the authors also conducted a sensitivity analysis in the paper for assessing the implications and limitations in the findings of the paper.

C. SENSITIVITY ANALYSIS

Many studies used several different data integrity techniques for managing data integrity in healthcare sector. It is a challenging task for any researcher or reader to identify the correct and accurate data integrity management approach for healthcare related challenges. The SLR provides a ranking analysis through fuzzy-AHP methodology for identifying and prioritizing data integrity techniques in healthcare. For assessing the difficulties and implications of the results, the authors have used the sample proportion method and found the average success proportion and confidence intervals on 95% for data integrity techniques used in healthcare sector previously. Table 19 represents the calculated data of proportion and confidence level. Firstly, for assessing the average proportion of techniques authors use equation (1).

$$P' = \frac{X}{n} \tag{1}$$

where P' is average proportion of specific technique, X is number of studies discussing that technique and n represents the total number of included studies in paper.

After analyzing the p' authors find the q' through following,

$$q' = (1 - p') \tag{2}$$

Here, q' represents a variable that holds the value of 1-p' for calculation.

TABLE 19. Summary of average proportion and their respective confidence level limit.

Technique	Number of Papers	Average proportion of techniques per study	Upper 95% Confidence Limit	Lower 95% Confidence Limit	Standard error
Blockchain	6	0.285	0.32	0.25	0.035
Authentication	1	0.047	0.066	0.028	0.019
Masked authenticated message extension	1	0.047	0.066	0.028	0.019
Secure Cloud	3	0.147	0.44	0.15	0.30
SWC- based secrete sharing	1	0.047	0.066	0.028	0.019
Cryptography	2	0.950	0.95	0.94	0.004
Markel-tree based Approach	1	0.047	0.066	0.028	0.019

TABLE 20. Recalculated figures after only showing the subjectively measured studies.

Technique	Number of Papers	Average proportion of techniques per study	Upper 95% Confidence Limit	Lower 95% Confidence Limit	Standard error
Blockchain	4	0.190	0.22	0.16	0.032
Authentication	1	0.047	0.066	0.028	0.019
Masked authenticated message extension	1	0.047	0.066	0.028	0.019
Secure Cloud	2	0.950	0.95	0.94	0.004
SWC- based secrete sharing	1	0.047	0.066	0.028	0.019
Cryptography	1	0.047	0.066	0.028	0.019
Markel-tree based Approach	1	0.047	0.066	0.028	0.019

Then for calculating the standard error authors use the equation (3).

$$E = z_c\sqrt{\frac{p'q'}{n}} \tag{3}$$

Here, E stands for the standard error and Z_c represents the Confidence level value that is predefined for different confidence levels (In our case its 1.96).

Now after calculating above equations, authors calculate the upper and lower confidence limit on 95% through the following function:

$$\text{Upper limit} = \mathrm{p}' + \mathrm{E} \tag{4}$$

$$\text{Lower limit} = \mathrm{p}' - \mathrm{E} \tag{5}$$

The above table describes the different techniques and their respective average proportion of success and confidence limit. For understanding the results better, the authors recalculate the figures after only extracting the studies measured subjectively in table 20.

As we conducted an analysis, results shown in table 19 & 20 describe that the difference in average blockchain proportion significantly different ($p' < 10$). Overall results show that average proportion of approximate of all techniques is lower when they are measured subjectively in comparison of all studies proportion in tables. However, the difficulty of calculating importance is affected by small number of studies that are not measured subjectively in the proposed paper.

VII. CONCLUSION

This SLR provides a brief knowledge of current scenario of data integrity in healthcare through attack statistics for better understanding. In fact, this investigation discusses prior studies in data integrity approaches to clarify the work situation to handle data integrity in the healthcare sector. The findings of this SLR show clearly that a modern and safer data integrity strategy is needed in the healthcare sector. The first section of this SLR shows the criticalness of data integrity issues in healthcare organizations. The second section (Review part) provides an idea for future researchers to adopt and motivate them for research in data integrity. With the help of priority assessment, this SLR may be helpful for learners. The ranking analysis assesses the priority of the data integrity techniques previously used in healthcare and ranks them through Fuzzy-AHP technique that would provide a path to the future researchers related to data integrity techniques and methods. Sensitivity analysis in this paper tries to find the difficulties and implications through statistical method.

The entire study was guided by two distinct objectives which have been detailed in the Discussion section. The first objective was to provide a brief and descriptive analysis of previous publications through different analysis methods. The second objective was to enumerate the data integrity techniques in all the methods discussed previously. Such a compilation would prove to be a depository for researchers and practitioners who are exploring both the possible

solutions to the problem of preserving data integrity as well as adopt the most prioritized technique to secure the data in the healthcare industry. The limitation of this study is the number of studies reviewed and databases that could be accessed by the authors. Although many databases were accessed by the authors, there are definitely some studies and databases which could not be incorporated in the profiled SLR.

REFERENCES

[1] R. Chakraborty, J. Mathew, and A. Vasilakos, Eds., "Security and fault tolerance in Internet of things," *Signal and Communication*. Springer, 2019, doi: 10.1007/978-3-030-02807-7.

[2] (2019). *Healthcare Data Breach Statistics*. Accessed: Oct. 21, 2019. [Online]. Available: https://www.hipaajournal.com/healthcare-data-breach-statistics/

[3] B. Filkins. (2014). *SANS Health Care Cyber-threat Report: Widespread Compromises Detected, Compliance Nightmare on Horizon. Norse*, Accessed: Oct. 21, 2019. [Online]. Available: https://www.sans.org/reading-room/whitepapers/firewalls/paper/34735

[4] (2018). *Breached Patient Records Tripled in 2018 vs 2017, as Health Data Security Challenges Worsen*. Accessed: Oct. 23, 2019. https://www.protenus.com/press/press-release/breached-patient-records-tripled-in-2018-vs-2017-as-health-data-security-challenges-worsen

[5] (2017), *Here's How Much Your Personal Information Is Selling for on the Dark Web*. Accessed: Oct. 23, 2019. https://www.experian.com/blogs/ask-experian/heres-how-much-your-personal-information-is-selling-for-on-the-dark-web/

[6] (2019). *Healthcare Data Breaches Reach Record High in April*. Accessed: Oct. 27, 2019. https://www.modernhealthcare.com/cybersecurity/healthcare-data-breaches-reach-record-high-april

[7] (2019), *The 10 Biggest Healthcare Data Breaches of 2019*. Accessed: Nov. 4, 2019. [Online]. Available: https://healthitsecurity.com/news/the-10-biggest-healthcare-data-breaches-of-2019-so-far

[8] "Data breach at major healthcare firms," *Comput. Fraud Secur.*, vol. 2019, no. 6, p. 3, Jun. 2019.

[9] (2019). *Healthcare Cybersecurity Market value to hit 27 billion by 2025 Global Market Insights*. Accessed: Nov. 4, 2019. [Online]. Available: https://www.globenewswire.com/news-release/2019/05/20/1827637/0/en/Healthcare-Cybersecurity-Market-value-to-hit-27-billion-by-2025-Global-Market-Insights-Inc.html

[10] (2019). *Healthcare Data Breaches Reach Record High April*. Accessed: Nov. 10, 2019. [Online]. Available: https://www.modernhealthcare.com/cybersecurity/healthcare-data-breaches-reach-record-high-april

[11] T. D. Oyetoyan, M. G. Jaatun, and D. S. Cruzes, "Measuring developers' software security skills, usage, and training needs," in *Exploring Security in Software Architecture and Design*, vol. 1. Hershey, PA, USA: IGI Global, 2019, doi: 10.4018/978-1-5225-6313-6.ch011.

[12] A. Reyes-Menendez, J. R. Saura, and F. Filipe, "The importance of behavioral data to identify online fake reviews for tourism businesses: A systematic review," *PeerJ Comput. Sci.*, vol. 5, p. e219, Sep. 2019, doi: 10.7717/peerj-cs.219.

[13] D. Moher, A. Liberati, J. Tetzlaff, and D. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *J. Clin. Epidemiol.*, vol. 62, no. 10, pp. 1006–1012, 2009, doi: 10.1016/j.jclinepi.2009.06.005.

[14] W. J. Gordon and C. Catalini, "Blockchain technology for healthcare: Facilitating the transition to patient-driven interoperability," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 224–230, 2018, doi: 10.1016/j.csbj.2018.06.003.

[15] J. Brogan, I. Baskaran, and N. Ramachandran, "Authenticating health activity data using distributed ledger technologies," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 257–266, 2018, doi: 10.1016/j.csbj.2018.06.004.

[16] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom, "FHIRChain: Applying blockchain to securely and scalably share clinical data," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 267–278, 2018, doi: 10.1016/j.csbj.2018.07.004.

[17] C. Esposito, A. De Santis, G. Tortora, H. Chang, and R. K.-K. Choo, "Blockchain: A panacea for healthcare cloud-based data security and privacy?" *IEEE Cloud Comput.*, vol. 5, no. 1, pp. 31–37, Jan./Feb. 2018, doi: 10.1109/mcc.2018.011791712.

[18] P. Gope and T. Hwang, "BSN-care: A secure IoT-based modern healthcare system using body sensor network," *IEEE Sensors J.*, vol. 16, no. 5, pp. 1368–1376, Mar. 2016, doi: 10.1109/jsen.2015.2502401.

[19] P. Vimalachandran, H. Wang, Y. Zhang, B. Heyward, and Y. Zhao, "Preserving patient-centred controls in electronic health record systems: A reliance-based model implication," in *Proc. Int. Conf. Orange Technol. (ICOT)*, Dec. 2017, pp. 37–44, doi: 10.1109/icot.2017.8336084.

[20] M. Elhoseny, G. Ramirez-Gonzalez, O. M. Abu-Elnasr, S. A. Shawkat, N. Arunkumar, and A. Farouk, "Secure medical data transmission model for IoT-based healthcare systems," *IEEE Access*, vol. 6, pp. 20596–20608, 2018, doi: 10.1109/access.2018.2817615.

[21] E. Luo, M. Z. A. Bhuiyan, G. Wang, M. A. Rahman, J. Wu, and M. Atiquzzaman, "PrivacyProtector: Privacy-protected patient data collection in IoT-based healthcare systems," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 163–168, Feb. 2018, doi: 10.1109/mcom.2018.1700364.

[22] M. Al Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *J. Med. Syst.*, vol. 36, no. 1, pp. 93–101, Mar. 2010, doi: 10.1007/s10916-010-9449-4.

[23] G. Manogaran, C. Thota, D. Lopez, and R. Sundarasekar, "Big data security intelligence for healthcare industry 4.0," *Cybersecur. Ind.*, vol. 4, pp. 103–126, Apr. 2017, doi: 10.1007/978-3-319-50660-9_5.

[24] B. Fabian, T. Ermakova, and P. Junghanns, "Collaborative and secure sharing of healthcare data in multi-clouds," *Inf. Syst.*, vol. 48, pp. 132–150, Mar. 2015, doi: 10.1016/j.is.2014.05.004.

[25] J. Sun, X. Zhu, C. Zhang, and Y. Fang, "HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare," in *Proc. 31st Int. Conf. Distrib. Comput. Syst.*, Jun. 2011, pp. 373–382, doi: 10.1109/icdcs.2011.83.

[26] A. Al Omar, S. M. Rahman, A. Basu, and S. Kiyomoto, "Medibchain: A blockchain based privacy preserving platform for healthcare data," in *Security, Privacy, and Anonymity in Computation, Communication, and Storage* (Lecture Notes in Computer Science). Springer, 2017, pp. 534–543, doi: 10.1007/978-3-319-72395-2_49.

[27] S. Bowman, "Impact of electronic health record systems on information integrity: Quality and safety implications," *Perspect. Health Inf. Manage.*, to be published.

[28] A. Theodouli, S. Arakliotis, K. Moschou, K. Votis, and D. Tzovaras, "On the design of a blockchain-based system to facilitate healthcare data sharing," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Communications/ 12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 1374–1379, doi: 10.1109/trustcom/bigdatase.2018.00190.

[29] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, "Integrating blockchain for data sharing and collaboration in mobile healthcare applications," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–5, doi: 10.1109/pimrc.2017.8292361.

[30] K. Abouelmehdi, A. Beni-Hessane, and H. Khaloufi, "Big healthcare data: Preserving security and privacy," *J. Big Data*, vol. 5, no. 1, Jan. 2018, doi: 10.1186/s40537-017-0110-7.

[31] A. Sajid and H. Abbas, "Data privacy in cloud-assisted healthcare systems: State of the art and future challenges," *J. Med. Syst.*, vol. 40, no. 6, May 2016, doi: 10.1007/s10916-016-0509-2.

[32] B. Sharma, C. N. Sekharan, and F. Zuo, "Merkle-tree based approach for ensuring integrity of electronic medical records," in *Proc. 9th IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Nov. 2018, pp. 983–987, doi: 10.1109/uemcon.2018.8796607.

[33] K. Gammon, "Experimenting with blockchain: Can one technology boost both data integrity and patients' pocketbooks?" *Nature Med.*, vol. 24, no. 4, pp. 378–381, 2018, doi: 10.1038/nm0418-378.

[34] D. J. Price. *Little Science, Big Science*. New York, NY, USA: Columbia Univ. Press, 1963, doi: 10.7312/pric91844.

[35] Z. Zheng, S. Xie, H. N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey," *Int. J. Web Grid Services*, vol. 14, no. 4, p. 352, 2018, doi: 10.1504/ijwgs.2018.095647.

[36] A Agrawal, M. Zarour, M. Alenezi, R. Kumar, RA. Khan, "Security durability assessment through fuzzy analytic hierarchy process," *PeerJ Comput. Sci.*, vol. 5, p. e215, Sep. 201, doi: 10.7717/peerj-cs.215.

[37] (2018). *Cloud in the Crosshairs*. Accessed: Jan. 15, 2020. [Online]. Available: https://www.netscout.com/report/

[38] *Advance Persistent Threat (APT)*, Accessed: Jan. 15, 2020. [Online]. Available: https://www.imperva.com/learn/application-security/apt-advanced-persistent-threat/

[39] *CVE Details*. Accessed: Jan. 17, 2020. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-26/product_id-251/Microsoft-Sql-Server.html

[40] AK. Pandey, AK. Tripathi, G. Kapil, V. Singh, M. W. Khan, A. Agrawal, and R. KumarRA. Khan, "Current Challenges of Digital Forensics in Cyber Security," in *Critical Concepts, Standards, and Techniques in Cyber Forensics*. Hershey, PA, USA: IGI Global, pp. 31–46, 2020, doi: 10.4018/978-1-7998-1558-7.ch003.
[41] A. Agrawal and N. R. Alharbe, "Need and importance of healthcare data integrity," *Int. J. Eng. Technol.*, vol. 11, no. 4, pp. 854–859, Aug. 2019.
[42] A. Pashazadeh and N. J. Navimipour, "Big data handling mechanisms in the healthcare applications: A comprehensive and systematic literature review," *J. Biomed. Informat.*, vol. 82, pp. 47–62, Jun. 2018.
[43] P. Pietro Biancone, S. Secinaro, V. Brescia, and D. Calandra, "Data quality methods and applications in health care system: A systematic literature review," *Int. J. Bus. Manage.*, vol. 14, no. 4, p. 35, Mar. 2019.
[44] H. M. Hussien, S. M. Yasin, S. N. I. Udzir, A. A. Zaidan, and B. B. Zaidan, "A systematic review for enabling of develop a blockchain technology in healthcare application: Taxonomy, substantially analysis, motivations, challenges, recommendations and future direction," *J. Med. Syst.*, vol. 43, no. 10, Sep. 2019.
[45] B. Pourghebleh and N. J. Navimipour, "Data aggregation mechanisms in the Internet of Things: A systematic review of the literature and recommendations for future research," *J. Netw. Comput. Appl.*, vol. 97, pp. 23–34, Nov. 2017.
[46] A. Vakili and N. J. Navimipour, "Comprehensive and systematic review of the service composition mechanisms in the cloud environments," *J. Netw. Comput. Appl.*, vol. 81, pp. 24–36, Mar. 2017.

ABHISHEK KUMAR PANDEY received the bachelor's degree (Hons.) in computer applications from Siddhartha University, Kapilvastu, India, in 2018. He is currently pursuing the master's degree in cyber security with Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, India. He is also a Passionate Researcher and frequently involved in various research initiatives. His area of interest in research is healthcare data security, malware analysis, digital forensic, and cyber security methods.

ASIF IRSHAD KHAN is currently working as an Assistant Professor with the Computer Science Department, King Abdulaziz University, Saudi Arabia. He has over 15 years of experience as a professional academician and a researcher. He has published several research articles in leading international journals and conferences. His current research interest includes software engineering with a focus on software security, component-based software engineering and software product line engineering.

YOOSEF B. ABUSHARK is currently an Assistant Professor with the Department of Computer Science, KAU. His research interests are in software engineering with a focus engineering intelligent systems and building agent-based simulations. He has been publishing several research outcomes in leading venues.

MD. MOTTAHIR ALAM received the Ph.D. degree in electronics and communication from Singhania University, India. He has six years of experience as a Software Engineer (quality) for leading software multinationals, where he worked on projects for companies like Pearson and Reader's Digest. He is currently an ISTQB Certified Software Tester and an Assistant Professor with the Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia. His research interests include software engineering (ESP) software product line engineering, software reusability, component-based and agent-based software engineering. He published several research articles in leading journals and conferences.

ALKA AGRAWAL received the Ph.D. degree from Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow. She is currently working as an Assistant Professor with the Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow. She is also a Passionate Researcher and has also published a number of research articles in national and international journals both. She has research/ teaching experience of more than 13 years. Her areas of research include software security and software vulnerability. She is also working in the fields of big data security, genetic algorithms, and software security.

RAJEEV KUMAR received the master's and Ph.D. degrees in information technology from Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, Uttar Pradesh, India, in 2014 and 2019, respectively. He is currently working as a Guest Faculty with the Department of Information Technology, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow. He is a young, energetic researcher and has worked on a Full Time Major Project funded by University Grants Commission, New Delhi, India. He has more than five years of research & teaching experience. He has also published & presented articles in refereed journals and conferences. His research interests are in the different areas of security engineering.

RAEES AHMAD KHAN is currently working as a Professor, the Head of the Department of the Department of Information Technology, and the Dean of the School for Information Science and Technology, Babasaheb Bhimrao Ambedkar University, (A Central University), Lucknow, India. He has more than 20 years of teaching & research experience. His areas of interest are software security, software quality, and software testing. He has published a number of National and International Books (including Chinese Language), Technical Article, Research Papers, Reviews and Chapters on Software Security, Software Quality, and Software Testing.