Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(REVIEW ARTICLE)

Check for updates

A survey on techniques, methods and security approaches in big data healthcare

David Odera *

Tom Mboya University, Faculty of biological and physical sciences, Homa-Bay, Kenya.

Global Journal of Engineering and Technology Advances, 2023, 14(02), 093–106

Publication history: Received on January 2023; revised on 18 February 2023; accepted on 21 February 2023

Article DOI: https://doi.org/10.30574/gjeta.2023.14.2.0035

Abstract

A huge percentage of people especially in developed countries spend a good chunk of their wealth in managing their health conditions. In order to adequately administer healthcare, governments and various organizations have embraced advanced technology for automating the health industry. In recent past, electronic health records have largely been managed by Enterprise Resource Planning and legacy systems. Big data framework steadily emerge as the underlying technology in healthcare, which offers solutions that limits capacity of others systems in terms of storage and reporting. Automation through cloud services supported by storage of structured and unstructured health data in heterogeneous environment has improved service delivery, efficiency, medication, diagnosis, reporting and storage in healthcare. The argument support the idea that big data healthcare still face information security concern, for instance patient image sharing, authentication of patient, botnet, correlation attacks, man-in-the-middle, Distributed Denial of Service (DDoS), blockchain payment gateway, time complexities of algorithms, despite numerous studies conducted by scholars in security management for big data in smart healthcare. Some security technique include digital image encryption, steganography, biometrics, rule-based policy, prescriptive analysis, blockchain contact tracing, cloud security, MapReduce, machine-learning algorithms, anonymizations among others. However, most of these security solutions and analysis performed on structured and semi-structured data as opposed to unstructured data. This may affect the output of medical reporting of patients' condition particularly on wearable devices and other examinations such as computerized tomography (CT) Scans among others. A major concern is how to identify inherent security vulnerabilities in big healthcare, which generate images for transmission and storage. Therefore, this paper conducted a comparative survey of solutions that specifically safeguards structured and unstructured data using systems that run on big data frameworks. The literature highlights several security advancements in cryptography, machine learning, anonymization and protocols. Most of these security frameworks lacks implementation evidence. A number of studies did not provide comprehensive performance metrics (accuracy, error, recall, precision) of the models besides using a single algorithm without validated justification. Therefore, a critique on the contribution, performance and areas of improvements discussed and summarized in the paper.

Keywords: HDFS; Blockchain; IoMH; MIT; WMSN; ANN; SVM; CNN; PUF; ECC

1. Introduction

There is continuous development of digital health solutions using modern technology in managing patients' treatments. The driving factor is sharing of health records among various agencies such as doctors, nurses, health administrators, pharmacies and hospitals. Authors in [1] note that autonomous databases from independent systems (silos) leads to poor integration of health records. In order to conduct proper analytics that can lead to correct insight in a patient medical [2] history these databases should converge in common pool to ease treatments, diagnosis, reporting and general utilization.

*Corresponding author: David Odera

The volume and granularities that comes with these data adds to the existing complexities [3]. It then creates a demand for integration and distribution of resources for both patients and providers in order to remedy medical cost and time. Internet of Things (IoT) [4], [5], cloud computing and big data frameworks are some of the known platforms used today for data capture, data processing and data storage in big healthcare [6]. The authors in [7] in proposing a new methodology to integrate big health and traditional business processes, singled out big data analytics as "a nascent field in health domain". The emergence of this platform presents barriers as well as potentials. In order to overcome them, there is a need to execute particular solutions for real situations on big health platform [8].

Big data in health is desirable [1]and technologically there is a common understanding by scientist on big data as been noted by [3]. The understanding of health practitioners may not mirror that of big data scientists. Hence, it leads to a knowledge gap. Perhaps health practitioners may be having a different perspective and comprehension of big data analytics. That begs the question why most health organizations are not making beneficial investment on big data with a good number being skeptical about it [3]. Security of their data including patients' data could be among the concerns. According to [9], tremendous violations of patients records where experienced in July 2019 where LabCorp Clinical Laboratory and Quest Diagnostics suffered breach that affected over 12 million people. In the same year, a report by IT health security in Canada, talks about 25 million sensitive patients' records were exposed leading to big loss economically. There is need to understand big data analytics and its application in economics, strategy, social, information security [11], [12], [13], [14] ,[15] [16] and professional issues in order to maximize the benefits that it brings. The usual security issues such as violation of access controls [17], alteration of clinical images, manipulative usage need to be contained [9]. These has led to application of techniques such as digital image encryption, steganography, biometrics, rule-based policy, prescriptive analysis, blockchain [18], [19], [20], [21], [22] contact tracing, cloud security, MapReduce, machine-learning algorithms, anonymizations [23] among others. Adequate big healthcare security framework is needed especially having experienced COVID19 patients' sensitive data [24], [25], [26] and medical images being shared through tele-diagnosis and online-consultations.

The objective of this study is to examine various studies in big healthcare security. The specific areas are as follows;

- To discuss meeting technologies in big data framework and its adoption in health
- Identifying security frameworks and solutions used in safeguarding information security in big data by showing their contributions, technologies, methods used and data sources
- Indicating research points and gaps and suggesting areas of further research work as far as big data security in health

2. The era of big data

The evolution of big data could be traced in 1997 when two data scientist (Michael Cox and David Ellsworth) presented the concept of data visualization and some of the challenges its facing [3]. Due to further innovations in databases, the concept of business intelligence followed which made it possible to discover knowledge in databases. The biggest concern was to enable online transaction processing in data warehouses. Application developments like Hadoop, web xml, database management systems included analytical modules and additional capabilities that enabled users to process big amount of data in real-time. It is at this time when hospitals and health organizations started to digitize their records [27], [28], [29] to facilitate storage, data manipulation, control, security, and aided healthcare to provide efficiency in data management.

The revolution stage boosted by big data pioneer industry in financial and commercial organizations, which used it in business processing, customer attraction and retention, workforce management in order to reduce the operational costs [30]. Today most analysis involves collection of data from cloud, IoT [31] and big data. The COVID 19 worldometer.info reported a huge crisis, which severely affected the health sectors around the world due to number of infections and fatality rate. The medical personnel were overstretched. It's one of the factors that drove numerous innovations in big digital health [32], [33], [34] such as internet of medical things [9], [35], blockchain-based eHealthcare [36], Secret Sharing Schemes [37], [38], image encryption and DNA coding [39], [40] among others.

2.1. Meeting technologies in Big Data

Internetworking of autonomous technologies is aiding implementation of big data frameworks. Big data concept is able to amass, integrate, process, store, analyze big datasets emanating within and out of organization [41]. The datasets may appear in structured, semi-structured and unstructured form. Therefore any dataset which exhibits the 5V's (Volume, Velocity, Variety, Veracity and Value) of big data in heterogeneity is considered big dataset [42]. Big data framework is supported by various meeting technologies such as parallel computing, Hadoop Distributed File System,

Map-Reduce. The author in [7] proposes a meeting technology that combines both traditional business intelligence (BI) and Big Data frameworks in order to handle transactional and analytical processing. To greater extent, BI is largely seen as part a parcel of big data framework with a functional difference. A Business Intelligence (BI) is defined as a system consisting of various technologies assisting organizations to collect, merge and analyze huge data to discover opportunities, strengths and weaknesses [43]. Similarities exist in storage of huge amount of data [44], knowledge discovery, integration, ad hoc reporting and forecasting [43]. Figure 1 below is an illustration of BI framework.

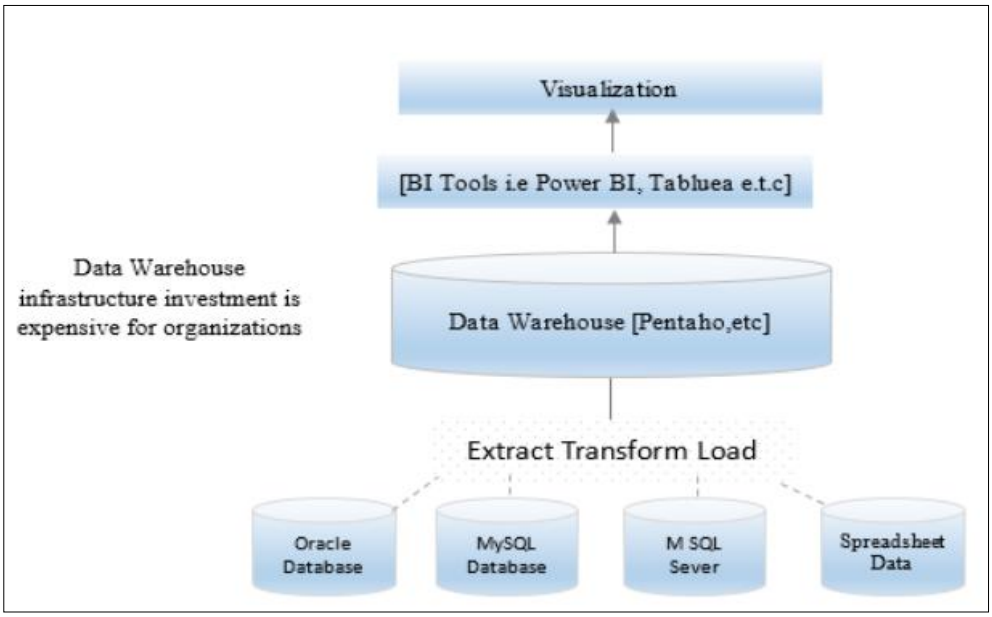

Figure 1 Business Intelligence (BI) Architecture

Big data framework consist of Hadoop that is a distributed processing environment and a MapReduce that provides programming model for batch processing. Hadoop distribute data by using primary node known as name node (stores meta data) and multiple data nodes which stores a replica of data for redundancy. The automatic data detection and restart of failed node in Hadoop manages fault tolerance through data replication. MapReduce is a functional programming model that splits data into map and reduce [45]. Apache Spark is another programming tool that can perform both real-time and batch processing of Hadoop file systems. Spark does well in processing due to short interval batches and streaming of API data. Following the discussion in [45], the Hadoop MapReduce framework consist of the following components;-

2.1.1. YARN

Is an abbreviation of Yet Another Resource Negotiator. Its major functions is to manage the resource [46] allocation to applications running on Hadoop. Some of the features it provides are memory organization, CPU management, tracking of running applications among others.

2.1.2. HDFS

Stands for Hadoop Distributed File System. It enables file replication so that failures of hardware is managed. It also supports optimization of processing because computation [47] is done on the location of file. Then linking of these data in several nodes are combine through same file system.

2.1.3. MAP-REDUCE

It provides the loci for mapping and reducing. Map function captures data from the file system and present the into pairs (Key, value) in their different sections. Reduce then collate the result from the mapper. The operations of Map and Reduce functions is divided into job submission node, Name node and Slave node. In every cluster, all data processing are done in Data node and Name node assumes the role of controlling all the Data nodes. Figure 2 below illustrates the big data architecture

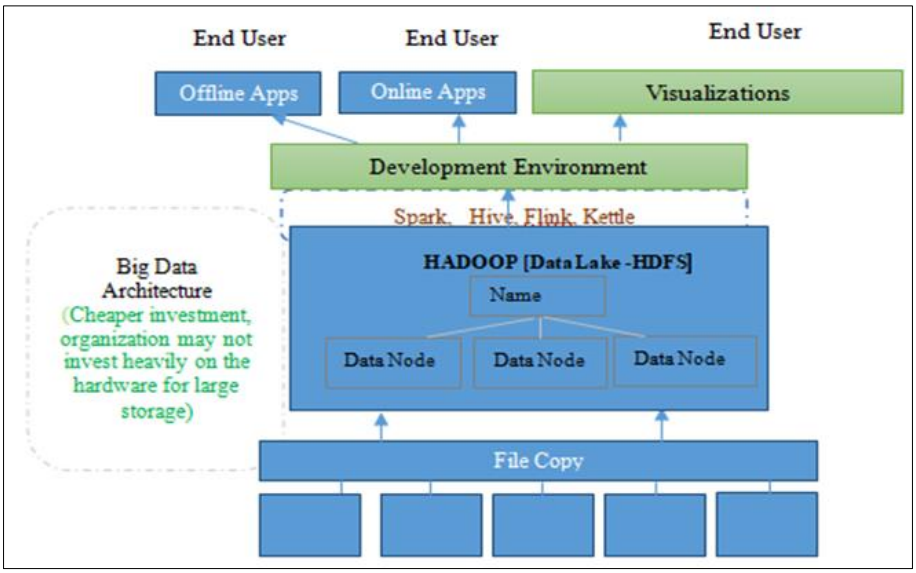

Figure 2 Hadoop Architecture

The recent advancement in Big Data framework is supported by programming platforms such as Apache Spark designed for real time systems. Spark's processing speed is faster in comparison with Hadoop because it uses in-memory computing [49] unlike disk-based computing used by Hadoop. It can support a number of programming languages such as python, java, R and Scala through high level APIs that enables graphs as well. It also supports machine learning through MLib, structured data management through Hive and graph management using graphX [50].

2.2. Big Healthcare Contributions

Solutions in big data healthcare have not only promoted sharing of electronic records between health entities but also treatment of patients online. The advancement of cloud services, big data, internet of medical things (IoMT) and the 5G communications technologies enables innovations such as wireless medical sensor networks (WMSN) possible [51].The authors in [52] proposes a similar but improved concept of a robust authentication protocol [53] for WMSN, which uses blockchain, and PUF in order to resolve security problems identified in [51]. Table 1 below outline some of the techniques, methods, algorithms, contributions and challenges that researchers have done in relation to big data healthcare.

Table 1 Big data healthcare

Technique	Contribution	Algorithm	Dataset	Challenge
A Personalized Ubiquitous Cloud and Edge-Enabled Networked Healthcare System for Smart Cities" in Big Data Learning & discovery by	UbeHealth overcome latency, bandwidth, reliability, security and energy efficiency issues. Disease co-occurrence designed using patient discharge database, mines data to predict future treatment	Deep learning, IoT	ISPDSL-II from 2013 and Waikato-VIII from 2018	Areas of improvement include feature security, privacy and more reliability
5G-Smart Diabetes Testbed	Personalized diabetes testbed *Validate performance using SVM, ANN & Decision tree*	SVM, ANN Decision tree		Metrix for performance validations such as Accuracy, Precision, Error, Recall are all not represented

Fuzzy rule based big data analytics-healthcare as a service [103]	Provided a cloud service to support healthcare in big data (health-as-a-service) using fuzzy logic *Reduce false positive, better accuracy*	Fuzzy rule classification	Dataset not clearly defined	Depends on the protection of platform infrastructure e.g cloud
Omic and Electronic Health Record Big Data Analytics for Precision Medicine [102]	Earlier precision disease prediction increases the efficiency of general	SVM	Multiomic data and EHR data. From the biological samples, molecular profiles(genomic, transcriptomic, epigenomic, proteomic, metabolomics	Need to explore performance validation using other algorithms
Emotion-Aware Connected Healthcare Big Data using Speech and image	Speech and video processing to identify patient medical status	SVM classifier	Health data visualization tool proposed	It's a framework Need to use other classifiers to predict
Diagnosis of non-small cell lung cancer using CNN	sensitivity of 85% to 90% in detecting false positive	CNN, back propagation algorithm, Hybrid support vector machine, K means, Deep learning, Supervised learning and fusion model	Ensemble methods lung cancer dataset	The sensitivity of detection level can be improved through optimization i.e reduction of error margin
Gamification is identified for predicting breast cancer	Efficient in medical image database accuracy of 81.3 - 82.7% in recurrence prediction Predicts NSCLC most affecting disease	CNN		Room for improvement on accuracy, security component is missing, adopt other models
voice pathology detection[104]	Accuracy of 97.5%	CNN algorithm and transfer learning	Saarbrucken voice disorder database.	Comparison with other algorithm such as RNN
Patient survival for lung cancer prediction using Unsupervised Deep Learning	Kaplan-Meier analysis and cox proportional hazards divide the patients groups into two, low risk and high risk accuracy has been improvised.	CNN convolutional auto encoder, LASSO-Cox model for feature selection		What of survival on recurring cases

2.3. Security in Big Data Healthcare

According to [54], information security in healthcare is linked with patient privacy issues. Security of information becomes even more important when you are developing health based information system [55], [56], [57], [58].

Table 2 Research on Big Data Security in Health

Task	Author	Contribution	Method	Gap
HealthCare EHR: A Blockchain-Based Decentralized Application	Panigrahi, Nayak & Rourab [54]	peer to peer network platform to build distributed database for data sharing among health entities	Ethereum blockchain	Online payment through the banks is not possible Adequate security is not guaranteed because all the participating peers in blockchain can access transactions
A Security Management Framework for Big Data in Smart Healthcare	Sarosh et al. [9]	The proposed scheme can be used to secure both natural as well as medical images. Reduces the requirement of resources like storage space and transmission bandwidth. generates small-sized shares for distributed storage of medical images in the IoT-based servers CSIS method-converts encrypted images into cloud-based servers Inverse CSIS-used to recover image.	logistic equation, hyperchaotic equation, and DNA encoding. hybrid chaos-based-image encryption	Initial problem was on more secure storage space of images and text However, this study does not address slower secret reconstruction process, computational complexity, and residual image problem.
Fog Computing Facility with Pairing-Based Cryptography	Al Hamid et al. [68]	Method to secure the MBD in the healthcare cloud by using DMBD, which depends on using a fog computing facility and pairing-based cryptography (PBC). A session key is generated for secure communication among the participants by using PBC to access and store MBD in the cloud.	Hash-based Message Authentication Code (HMAC) placed on the decoy document. In the event of the user accessing only the DMBD, an SMS or email will be sent to the legitimate user to inform him/her that his/her account has been accessed. hybrid User Profiling Algorithm Elliptic Curve Cryptography was selected to maintain Key exchange can happen by either RSA or ECC algorithms.	Communication Overhead User, OMBD, and DMBD A tri-party authenticated key agreement protocol has been proposed among the user for efficiency but not designed and implemented Evaluation of technique is not done

An End-to-End Security Framework for Smart Healthcare Information Sharing against Botnet-based Cyber-Attacks	Quamara et al. [69]	Highlight healthcare infrastructure and components Used simple rule based policy using encryption key	Rule-based policy framework. Access Control Policy Testing (ACPT) tool developed by National Institute of Standard (NIST)	Tried to solve Botnet-based cyber-attacks The framework is in development and hence, the perspective work would be in real-time implementation of the same and extension of the botnet detection mechanisms towards protocol the security policies
Blockchain Meets COVID-19: A Framework for Contact Information Sharing and Risk Notification System	Song et al. [70]	Bluetooth technology	blockchain	Tried to solve Contact tracing of potential victims however, Time complexity of the framework [71] is good, however it may affect performance[72] need to optimize, validate
Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations	Wang et al. [73]	Identify big data analytics capabilities to explore the potential benefit it brings content analysis of 26 big data implementation cases in health care	Quantitative approach Content Analysis that is empirically grounded method, its described as exploratory in process and predictive in intent	Suggestion to use machine learning [74], [75], [76], [77], [78] algorithm in analyzing unstructured data is indicated as future research
Proposed Meta Cloud-Redirection (MC-R) architecture with big data knowledge system is used to collect and store the sensor data (big data)	Manogaran et al. [79]	Provide threat and attacks against IoT devices and Available security requirements and solutions Cloud computing storage using Amazon Classification of data in database as sensitive, normal and critical Logging of users access to the database	Uses Iaas for Storage, MapReduce	data synchronization is needed to integrate Cloud and IoT healthcare system [80], [81], [82], [83], [84] transmission overhead(remove multicast/broadcast). medical information rather than conventional centralized data storage to support fault tolerance Available security requirements [85] and solutions are not descriptive towards threat, attacks
Big healthcare data: preserving security and privacy	Abouelmehdi et al. [86]	UNC Health Care (UNCHC), clinicians analyze unstructured patient data using natural-language	a big data security event monitoring system model-spark Data collection(logs) Data integration(filtering, classifying) data analysis(correlation, association)	How to reconcile quality of data when using privacy preserving techniques [87], [88], [89], [90], [91] in big data is a concern

			data interpretation(visuals, statistical outputs)	
Making big data, privacy, and anonymization work together in the enterprise: experiences and issues. In: Big data	Sedayao et al. [92]	Contribution in data masking of sensitive part of private information	K-anonymity	it still uses K-anonymity technique which is vulnerable to correlation attack
Hiding a needle in a haystack: privacy preserving Apriori algorithm in MapReduce framework	Jung et al. [93]	analyze the anonymized data and acquire valuable results	k-anonymity based metrics, used MapReduce to execute and analyze the anonymized data and acquire valuable results	Solve privacy violation [93], [94], [95] without utility degradation but its execution time is affected by noise size
Disease Prediction by Machine Learning over Big Data from Healthcare Communities	Chen et al. [96]	new convolutional neural network based multimodal disease risk prediction (CNN-MDRP) -combine the structured and unstructured data in healthcare field to assess the risk of disease.	- CNN-based multimodal disease risk prediction (CNN-MDRP) algorithm -latent factor model to reconstruct (missing data) - statistical knowledge(determine the major chronic diseases in the region)	-Analysis of graphical data which nowadays is greatly being generated by most healthcare devices i.e telemedicine Effect of reconstruction of missing data on the validity of accuracy
Healthcare Biometrics Security and Regulations: Biometrics Data Security and Regulations Governing PHI and HIPAA Act for Patient Privacy	Jayanthilladevi et al. [97]	Enforcing HIPPA act using biometric technique to preserve confidentiality and privacy of patient information	Policy framework	No evidence of technical demonstration of biometric solution
Internet of Thing Based HealthCare Monitoring System	Chaudhury et al. [98]	We reviewed the IOT based technologies [99] being used as smart hospitals in present time, their mechanisms, advantages and disadvantage	Developed design principle of working of telehealth	designing systems still more advanced to put through the existing drawbacks(implementation) no implementation
Methodology for Big Data Analytics technologies into traditional business intelligence	Gonzalez-Alonso Vilar & Villanueva [7]	Integration of technologies of BI and Big data [100], [101], [102]	Framework approach	The approach lacks implementation, however it's also similar to cloud computing concept

The technology of blockchain which has driven exchange of crypto-currency particularly bitconin 2009, now forms the basis of other research solutions in healthcare. Blockchain is used to develop distributed database also refer to as a ledger, which is available for all participating nodes in a peer-to-peer network [59]. Ethereum blockchain platform has been used in [54] to build Electronic Health Record (EHR) and distributed application [60] which enable secure sharing [61] of records between doctors, patients, insurance companies. The challenge with this concept is integration with financial payment systems. Moreover, the concept of open ledger may be exploited as a vulnerability against distributed database.

Contribution such as usage of Radio Frequency Identification (RFID) to locate and authentic the right person in tele-Medicare [62], [63], [64] information technology is proposed in literature. RFID operates by help of radio waves to capture and store data on the tag and then use tagged details for authentication. It has enabled tele-medicare Information Technology (TMIT) particularly in verification process. Authors in [51] propose a Proerif verification tool and security attribute analysis for TMIT. It uses PUF Physical Unclonable Function (PUF) and Elliptic Curve Cryptography (ECC). The PUF provides secure authentication [65], [66], [67] between the server and the tag, while ECC does the encryption of PUF information. The result is a lightweight authentication scheme. The challenge with PUF technology is poor stability and hardware aging. Therefore, their future work is on how to improve quality of PUF technology. Table 2 below contains security frameworks and techniques that have been proposed and some tested by various researchers in big healthcare security.

3. Conclusion

Big data in health is critically important in sharing of health records and in analysis of patient's data. Most health organizations have adopted this framework in phased-based approach. The revolution of big data in health is seen during COVID19 pandemic where so many innovations (IoMT, eHealthcare, image encryption and DNA coding, Contact tracing using bluetooth) running on IoT, blockchain, cloud and big data were implemented. An attempt by researchers to integrate old analytical frameworks such as traditional business intelligence (BI) and systems that run on Hadoop Distributed File Systems have been proposed. Information security is point of concern because vulnerabilities, threats and weaknesses still exist in the frameworks. The incidences reported in LabCorp Clinical Laboratory and Quest Diagnostics proves that attackers can still exploit these weaknesses. Security advancement seen in areas such as authentication of patient using PUF technology, use of machine learning to detect and predict diseases, privacy preservation using k-anonymization, secure secret sharing, image encryption, paring-based cryptography among others, is commendable. Medical reporting of critical patients' condition depends on wearable devices for example machines used in conducting CT-Scans, MRI, radiology etc. Challenges of Security, storage and performance of these wearable devices in relation to big healthcare should be addressed effectively. While appreciating the major contributions of research studies in big healthcare, this paper also suggests areas of improvement as summarized in the tables above. According to literature in this paper, most models implement single machine learning algorithm, which result into development of simulations for analysis. Most of these simulations are still under development and therefore future research should focus on validity testing of models in order to improve reliability of the models in big data security analysis in health. A study also needs to be conducted to address susceptibility of correlation attack in K-anonymity technique, which is used in masking of sensitive part of private information.

Compliance with ethical standards

Acknowledgments

I would like to extend my gratitude to all my colleagues who offered a helping hand during the development of this manuscript.

References

[1] Panda M, Ali SM, Panda SK. Big data in health care: A mobile based solution. In2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC) 2017 Mar 23 (pp. 149-152). IEEE.

[2] Al-Shaher MA, Al-Khafaji NJ. E-healthcare system to monitor vital signs. In2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) 2017 Jun 29 (pp. 1-5). IEEE.

[3] Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological forecasting and social change. 2018 Jan 1, 126:3-13.

[4] Ahmad T, Ranise S. Validating Requirements of Access Control for Cloud-Edge IoT Solutions (Short Paper). InFoundations and Practice of Security: 11th International Symposium, FPS 2018, Montreal, QC, Canada, November 13–15, 2018, Revised Selected Papers 11 2019 (pp. 131-139). Springer International Publishing.

[5] Nyangaresi VO. Privacy Preserving Three-factor Authentication Protocol for Secure Message Forwarding in Wireless Body Area Networks. Ad Hoc Networks. 2023 Feb 8:103117.

[6] Al-Shaher MA, Al-Khafaji NJ. E-healthcare system to monitor vital signs. In2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) 2017 Jun 29 (pp. 1-5). IEEE.

[7] Gonzalez-Alonso P, Vilar R, Lupiáñez-Villanueva F. Meeting technology and methodology into health big data analytics scenarios. In2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS) 2017 Jun 22 (pp. 284-285). IEEE.

[8] Al-Zubaidie M, Zhang Z, Zhang J. PAX: Using pseudonymization and anonymization to protect patients' identities and data in the healthcare system. International Journal of Environmental Research and Public Health. 2019 May, 16(9):1490.

[9] Sarosh P, Parah SA, Bhat GM, Muhammad K. A security management framework for big data in smart healthcare. Big Data Research. 2021 Jul 15,25:100225.

[10] Vimalachandran P, Wang H, Zhang Y, Zhuo G. The Australian PCEHR system: ensuring privacy and security through an improved access control mechanism. arXiv preprint arXiv:1710.07778. 2017 Oct 21.

[11] Nyangaresi VO, Ogundoyin SO. Certificate based authentication scheme for smart homes. In2021 3rd Global Power, Energy and Communication Conference (GPECOM) 2021 Oct 5 (pp. 202-207). IEEE.

[12] Parviainen P, Tihinen M, Kääriäinen J, Teppola S. Tackling the digitalization challenge: how to benefit from digitalization in practice. International journal of information systems and project management. 2017, 5(1):63-77.

[13] Pilares IC, Azam S, Akbulut S, Jonkman M, Shanmugam B. Addressing the challenges of electronic health records using blockchain and ipfs. Sensors. 2022 May 26, 22(11):4032.

[14] Edmunds M, Peddicord D, Frisse ME. Ten reasons why interoperability is difficult. Healthcare information management systems: Cases, strategies, and solutions. 2016:127-37.

[15] Sahi MA, Abbas H, Saleem K, Yang X, Derhab A, Orgun MA, Iqbal W, Rashid I, Yaseen A. Privacy preservation in e-healthcare environments: State of the art and future directions. Ieee Access. 2017 Oct 30, 6:464-78.

[16] Kumar JS, Patel DR. A survey on internet of things: Security and privacy issues. International Journal of Computer Applications. 2014 Jan 1, 90(11).

[17] Nyangaresi VO. Lightweight anonymous authentication protocol for resource-constrained smart home devices based on elliptic curve cryptography. Journal of Systems Architecture. 2022 Dec 1, 133:102763.

[18] Azbeg K, Ouchetto O, Andaloussi SJ. Access Control and Privacy-Preserving Blockchain-Based System for Diseases Management. IEEE Transactions on Computational Social Systems. 2022 Jul 8.

[19] Zhang D, Wang S, Zhang Y, Zhang Q, Zhang Y. A secure and privacy-preserving medical data sharing via consortium blockchain. Security and Communication Networks. 2022 May 18, 2022.

[20] Ahmad RW, Salah K, Jayaraman R, Yaqoob I, Ellahham S, Omar M. The role of blockchain technology in telehealth and telemedicine. International journal of medical informatics. 2021 Apr 1, 148:104399.

[21] Mackey TK, Kuo TT, Gummadi B, Clauson KA, Church G, Grishin D, Obbad K, Barkovich R, Palombini M. 'Fit-for-purpose?'–challenges and opportunities for applications of blockchain technology in the future of healthcare. BMC medicine. 2019 Dec, 17(1):1-7.

[22] Wang S, Zhang Y, Zhang Y. A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems. Ieee Access. 2018 Jun 29, 6:38437-50.

[23] Nyangaresi VO. Provably Secure Pseudonyms based Authentication Protocol for Wearable Ubiquitous Computing Environment. In2022 International Conference on Inventive Computation Technologies (ICICT) 2022 Jul 20 (pp. 1-6). IEEE.

[24] Demir O, Kocak B. A decentralized file sharing framework for sensitive data. InBig Data Innovations and Applications: 5th International Conference, Innovate-Data 2019, Istanbul, Turkey, August 26–28, 2019, Proceedings 5 2019 (pp. 142-149). Springer International Publishing.

[25] Xu J, Gao X, Sorwar G, Croll P. Implementation of e-health record systems in Australia. The International Technology Management Review. 2013 Jul, 3(2):92-104.

[26] Shahid J, Ahmad R, Kiani AK, Ahmad T, Saeed S, Almuhaideb AM. Data protection and privacy of the internet of healthcare things (IoHTs). Applied Sciences. 2022 Feb 12, 12(4):1927.

[27] Porambage P, Ylianttila M, Schmitt C, Kumar P, Gurtov A, Vasilakos AV. The quest for privacy in the internet of things. IEEE Cloud Computing. 2016 May 25, 3(2):36-45.

[28] Hossain CA, Mohamed MA, Zishan MS, Ahsan R, Sharun SM. Awareness on E-Health among undergraduate students in Bangladesh. Indian J Public Health Res Dev. 2019 Mar 1, 10:636-41.

[29] Badr S, Gomaa I, Abd-Elrahman E. Multi-tier blockchain framework for IoT-EHRs systems. Procedia Computer Science. 2018 Jan 1, 141:159-66.

[30] Institute for Health Technology Transformation, 2013. Transforming Health Care through Big Data: Strategies for Leveraging Big Data in the Health Care Industry. Institute for Health Technology Transformation, New York.

[31] Nyangaresi VO, Ma J. A Formally Verified Message Validation Protocol for Intelligent IoT E-Health Systems. In2022 IEEE World Conference on Applied Intelligence and Computing (AIC) 2022 Jun 17 (pp. 416-422). IEEE.

[32] Alansari Z, Soomro S, Belgaum MR, Shamshirband S. The rise of Internet of Things (IoT) in big healthcare data: review and open research issues. Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2016, Volume 2. 2018:675-85.

[33] Zheng Z, Xie S, Dai H, Chen X, Wang H. An overview of blockchain technology: Architecture, consensus, and future trends. In2017 IEEE international congress on big data (BigData congress) 2017 Jun 25 (pp. 557-564). Ieee.

[34] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health information science and systems. 2014 Dec, 2:1-0.

[35] Amaraweera SP, Halgamuge MN. Internet of things in the healthcare sector: overview of security and privacy issues. Security, privacy and trust in the IoT environment. 2019:153-79.

[36] Wang W, Sun L, Liu T, Lai T. The use of E-health during the COVID-19 pandemic: a case study in China's Hubei province. Health Sociology Review. 2022 Sep 2, 31(3):215-31.

[37] Nyangaresi VO. Masked Symmetric Key Encrypted Verification Codes for Secure Authentication in Smart Grid Networks. In 2022 4th Global Power, Energy and Communication Conference (GPECOM) 2022 Jun 14 (pp. 427-432).

[38] Yan X, Lu Y, Liu L, Song X. Reversible image secret sharing. IEEE Transactions on Information Forensics and Security. 2020 Jun 11, 15:3848-58.

[39] Lu Q, Zhu C, Deng X, An efficient image encryption scheme based on the LSS chaotic map and single S-box, IEEE Access 8 (2020) 25664–25678.

[40] Wan Y, Gu S, Du B, A new image encryption algorithm based on composite chaos and hyperchaos combined with DNA coding, Entropy 22 (2) (2020) 171

[41] Akutota T, Choudhury S. Big data security challenges: An overview and application of user behavior analytics. Int. Res. J. Eng. Technol. 2017, 4:1544-8.

[42] Azaria A, Ekblaw A, Vieira T, Lippman A. Medrec: Using blockchain for medical data access and permission management. In2016 2nd international conference on open and big data (OBD) 2016 Aug 22 (pp. 25-30). IEEE.

[43] Ul-Ain N, Vaia G, DeLone W. Business intelligence system adoption, utilization and success-A systematic literature review.

[44] Nyangaresi VO. Lightweight key agreement and authentication protocol for smart homes. In2021 IEEE AFRICON 2021 Sep 13 (pp. 1-6). IEEE.

[45] Sogodekar M, Pandey S, Tupkari I, Manekar A. Big data analytics: hadoop and tools, 2016 IEEE Bombay Section Symposium (IBSS), 21-22 December 2016.

[46] Tao F, Zuo Y, Da Xu L, Zhang L. IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing. IEEE transactions on industrial informatics. 2014 Feb 17, 10(2):1547-57.

[47] Gordon WJ, Catalini C. Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. Computational and structural biotechnology journal. 2018 Jan 1, 16:224-30.

[48] Mamoshina P, Ojomoko L, Yanovich Y, Ostrovski A, Botezatu A, Prikhodko P, Izumchenko E, Aliper A, Romantsov K, Zhebrak A, Ogu IO. Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. Oncotarget. 2018 Jan 1, 9(5):5665.

[49] Alsamhi SH, Shvetsov AV, Kumar S, Shvetsova SV, Alhartomi MA, Hawbani A, Rajput NS, Srivastava S, Saif A, Nyangaresi VO. UAV computing-assisted search and rescue mission framework for disaster and harsh environment mitigation. Drones. 2022 Jun 22, 6(7):154.

[50] Salloum S, Dautov R, Chen X, Peng PX, Huang JZ. Big data analytics on Apache Spark. International Journal of Data Science and Analytics. 2016 Nov, 1:145-64.

[51] Wang W, Chen Q, Yin Z, Srivastava G, Gadekallu TR, Alsolami F, Su C. Blockchain and PUF-based lightweight authentication protocol for wireless medical sensor networks. IEEE Internet of Things Journal. 2021 Oct 5, 9(11):8883-91.

[52] Yu J, Park J, Hyun SS. Impacts of the COVID-19 pandemic on employees' work stress, well-being, mental health, organizational citizenship behavior, and employee-customer identification. Journal of Hospitality Marketing & Management. 2021 Jul 4, 30(5):529-48.

[53] Nyangaresi VO. Hardware assisted protocol for attacks prevention in ad hoc networks. InEmerging Technologies in Computing: 4th EAI/IAER International Conference, iCETiC 2021, Virtual Event, August 18–19, 2021, Proceedings 4 2021 (pp. 3-20). Springer International Publishing.

[54] Justinia T. Saudi Arabia: Transforming Healthcare with Technology. InNursing Informatics: A Health Informatics, Interprofessional and Global Perspective 2022 Jul 26 (pp. 755-769). Cham: Springer International Publishing.

[55] Hasselgren A, Kralevska K, Gligoroski D, Pedersen SA, Faxvaag A. Blockchain in healthcare and health sciences—A scoping review. International Journal of Medical Informatics. 2020 Feb 1, 134:104040.

[56] Zou R, Lv X, Zhao J. SPChain: Blockchain-based medical data sharing and privacy-preserving eHealth system. Information Processing & Management. 2021 Jul 1, 58(4):102604.

[57] Thwin TT, Vasupongayya S. Blockchain-based access control model to preserve privacy for personal health record systems. Security and Communication Networks. 2019 Jun 25, 2019.

[58] Chenthara S, Ahmed K, Wang H, Whittaker F, Chen Z. Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology. Plos one. 2020 Dec 9, 15(12):e0243043.

[59] Pop RA, Săplăcan Z, Dabija DC, Alt MA. The impact of social media influencers on travel decisions: The role of trust in consumer decision journey. Current Issues in Tourism. 2022 Mar 4, 25(5):823-43.

[60] Yang X, Li T, Xi W, Chen A, Wang C. A blockchain-assisted verifiable outsourced attribute-based signcryption scheme for EHRs sharing in the cloud. IEEE Access. 2020 Sep 18, 8:170713-31.

[61] Nyangaresi VO. A Formally Validated Authentication Algorithm for Secure Message Forwarding in Smart Home Networks. SN Computer Science. 2022 Jul 9, 3(5):364.

[62] Buldakova TI, Sokolova AV. Network services for interaction of the telemedicine system users. In2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA) 2019 Nov 20 (pp. 387-391). IEEE.

[63] Bashshur RL, Shannon G, Krupinski EA, Grigsby J. Sustaining and realizing the promise of telemedicine. Telemedicine and e-Health. 2013 May 1, 19(5):339-45.

[64] Andrès E, Talha S, Benyahia AA, Keller O, Hajjam M, Moukadem A, Dieterlen A, Hajjam J, Ervé S, Hajjam A. E-health: A promising solution for optimizing management of chronic diseases. Example of the national e-health project e-care based on an e-platform in the context of chronic heart failure. European Research in Telemedicine/La Recherche Européenne en Télémédecine. 2015 Sep 1, 4(3):87-94.

[65] Kheshaifaty N, Gutub A. Engineering graphical captcha and AES crypto hash functions for secure online authentication. Journal of Engineering Research. 2021 Nov 10.

[66] Ni J, Lin X, Shen XS. Efficient and secure service-oriented authentication supporting network slicing for 5G-enabled IoT. IEEE Journal on Selected Areas in Communications. 2018 Mar 12, 36(3):644-57.

[67] Nyangaresi VO, Ahmad M, Alkhayyat A, Feng W. Artificial neural network and symmetric key cryptography based verification protocol for 5G enabled Internet of Things. Expert Systems. 2022 Dec, 39(10):e13126.

[68] Al Hamid HA, Rahman SM, Hossain MS, Almogren A, Alamri A. 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. 2017 Sep 28, 5:22313-28.

[69] Quamara M, Gupta BB, Yamaguchi S. An end-to-end security framework for smart healthcare information sharing against botnet-based cyber-attacks. In2021 IEEE International Conference on Consumer Electronics (ICCE) 2021 Jan 10 (pp. 1-4). IEEE.

[70] Song J, Gu T, Fang Z, Feng X, Ge Y, Fu H, Hu P, Mohapatra P. Blockchain meets COVID-19: A framework for contact information sharing and risk notification system. In2021 IEEE 18th international conference on mobile ad hoc and smart systems (MASS) 2021 Oct 4 (pp. 269-277). IEEE.

[71] Stamatellis C, Papadopoulos P, Pitropakis N, Katsikas S, Buchanan WJ. A privacy-preserving healthcare framework using hyperledger fabric. Sensors. 2020 Nov 18, 20(22):6587.

[72] Nyangaresi VO. Terminal independent security token derivation scheme for ultra-dense IoT networks. Array. 2022 Sep 1, 15:100210.

[73] Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological forecasting and social change. 2018 Jan 1, 126:3-13.

[74] Li, J., Zhou, Z., Dong, J., et al., 2021. Predicting breast cancer 5-year survival using machine learning: A systematic review. PloS one. 16(4), e0250370.

[75] Binder, A., Bockmayr, M., Hägele, M., et al., 2021. Morphological and molecular breast cancer profiling through explainable machine learning. Nature Machine Intelligence. 3(4), 355-366.

[76] Haque, M.N., Tazin, T., Khan, M.M., et al., 2022. Predicting Characteristics Associated with Breast Cancer Survival Using Multiple Machine Learning Approaches. Computational and Mathematical Methods in Medicine.

[77] Shravya, C., Pravalika, K., Subhani, S., 2019. Prediction of breast cancer using supervised machine learning techniques. International Journal of Innovative Technology and Exploring Engineering (IJITEE). 8(6), 1106-1110.

[78] Nyangaresi, V.O., El-Omari, N.K.T., Nyakina, J.N., 2022. Efficient Feature Selection and ML Algorithm for Accurate Diagnostics. Journal of Computer Science Research. 4(1), 10-19.

[79] Manogaran G, Thota C, Lopez D, Sundarasekar R. Big data security intelligence for healthcare industry 4.0. Cybersecurity for Industry 4.0: Analysis for Design and Manufacturing. 2017:103-26.

[80] Akhbarifar S, Javadi HH, Rahmani AM, Hosseinzadeh M. A secure remote health monitoring model for early disease diagnosis in cloud-based IoT environment. Personal and Ubiquitous Computing. 2020 Nov 16:1-7.

[81] Hammi B, Khatoun R, Zeadally S, Fayad A, Khoukhi L. Internet of things (iot) technologies for smart cities, IET Networks 7, 2017.

[82] Hwang YH. Iot security & privacy: threats and challenges. InProceedings of the 1st ACM workshop on IoT privacy, trust, and security 2015 Apr 14 (pp. 1-1).

[83] Sadek I, Rehman SU, Codjo J, Abdulrazak B. Privacy and security of IoT based healthcare systems: concerns, solutions, and recommendations. InHow AI Impacts Urban Living and Public Health: 17th International Conference, ICOST 2019, New York City, NY, USA, October 14-16, 2019, Proceedings 17 2019 (pp. 3-17). Springer International Publishing.

[84] Chung, B.,Kim, J.,Jeon, Y. On-demand security configuration for IoT devices. In Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 19–21 October 2016,pp. 1082–1084.

[85] Nyangaresi VO. A Formally Verified Authentication Scheme for mmWave Heterogeneous Networks. Inthe 6th International Conference on Combinatorics, Cryptography, Computer Science and Computation (605-612) 2021.

[86] Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. Journal of big data. 2018 Dec, 5(1):1-8.

[87] Sahama T, Simpson L, Lane B. Security and Privacy in eHealth: Is it possible?. In2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013) 2013 Oct 9 (pp. 249-253). IEEE.

[88] Grobauer B, Walloschek T, Stocker E. Understanding cloud computing vulnerabilities. IEEE Security & privacy. 2010 Jun 17, 9(2):50-7.

[89] Rayes A, Salam S, Dabbagh M, Rayes A. Internet of things security and privacy. Internet of Things From Hype to Reality: The Road to Digitization. 2017:195-223.

[90] De Filippi P. The interplay between decentralization and privacy: the case of blockchain technologies. Journal of Peer Production, Issue. 2016 Sep 14(7).

[91] Nyangaresi VO, Mohammad Z. Session Key Agreement Protocol for Secure D2D Communication. InThe Fifth International Conference on Safety and Security with IoT: SaSeIoT 2021 2022 Jun 12 (pp. 81-99). Cham: Springer International Publishing.

[92] Sedayao J, Bhardwaj R, Gorade N. Making big data, privacy, and anonymization work together in the enterprise: experiences and issues. In2014 IEEE International Congress on Big Data 2014 Jun 27 (pp. 601-607). IEEE.

[93] Jung K, Park S, Park S. Hiding a needle in a haystack: privacy preserving apriori algorithm inmapreduce framework. InProceedings of the First International Workshop on Privacy and Secuirty of Big Data 2014 Nov 7 (pp. 11-17).

[94] Abdullah S, Arshad J, Khan MM, Alazab M, Salah K. PRISED tangle: a privacy-aware framework for smart healthcare data sharing using IOTA tangle. Complex & Intelligent Systems. 2022 Jan 21:1-9.

[95] Pussewalage HS, Oleshchuk VA. Privacy preserving mechanisms for enforcing security and privacy requirements in E-health solutions. International Journal of Information Management. 2016 Dec 1, 36(6):1161-73.

[96] Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big data from healthcare communities. Ieee Access. 2017 Apr 26, 5:8869-79.

[97] Jayanthilladevi A, Sangeetha K, Balamurugan E. Healthcare Biometrics Security and Regulations: Biometrics Data Security and Regulations Governing PHI and HIPAA Act for Patient Privacy. In2020 International Conference on Emerging Smart Computing and Informatics (ESCI) 2020 Mar 12 (pp. 244-247). IEEE.

[98] Chaudhury S, Paul D, Mukherjee R, Haldar S. Internet of Thing based healthcare monitoring system. In2017 8th annual industrial automation and electromechanical engineering conference (IEMECON) 2017 Aug 16 (pp. 346-349). IEEE.

[99] Samih H. Smart cities and internet of things. Journal of Information Technology Case and Application Research. 2019 Jan 2, 21(1):3-12.

[100] Chen M, Yang J, Zhou J, Hao Y, Zhang J, Youn CH. 5G-smart diabetes: Toward personalized diabetes diagnosis with healthcare big data clouds. IEEE Communications Magazine. 2018 Apr 13, 56(4):16-23.

[101] Hossain MS, Muhammad G. Emotion-aware connected healthcare big data towards 5G. IEEE Internet of Things Journal. 2017 Nov 13, 5(4):2399-406.

[102] Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. –Omic and electronic health record big data analytics for precision medicine. IEEE Transactions on Biomedical Engineering. 2016 Oct 10, 64(2):263-73.

[103] Jindal A, Dua A, Kumar N, Vasilakos AV, Rodrigues JJ. An efficient fuzzy rule-based big data analytics scheme for providing healthcare-as-a-service. In2017 IEEE international conference on communications (ICC) 2017 May 21 (pp. 1-6). IEEE.

[104] Alhussein M, Muhammad G. Voice pathology detection using deep learning on mobile healthcare framework. IEEE Access. 2018 Jul 16, 6:41034-41.