A Survey On Big Data Analytics In Medical And Healthcare Using Cloud Computing

Sreehari Kundella, Dr. R. Gobinath

Abstract: Big data analytics is the most emerging technologies in the world. On the previous few years, big data analytics has showed up as being a fresh paradigm of abundant information and possibilities for improving and research that is allowing big data. The way organizations handle, evaluate and leverage information in any sector has essentially altered. Big Data Analytics will be provide cloud-based data management and data management alternatives to analyze, store and process enormous data volumes. One of the most areas that are guaranteeing big data analytics used to come up with a change in medical care. Big data analytics in Medical and healthcare has important prospect of improving client outcomes, predicting epidemic outbreaks, gaining helpful insights, avoiding diseases, reducing health care distribution expenses, and increasing overall total wellbeing. You will find huge amounts of information in health care as well as a focus that is extra healthcare. Healthcare and medical information is increasing and increasing, many complicated, and their sources have actually increased to incorporate physician/doctor that is computerized entry, EMR, notes which are related to clinical, images, genomic information, and support systems in respect of health care and medical. In healthcare big data analytics involves integrating and analyzing huge degrees of complicated information that is heterogeneous as multiple omics information, medical information and electronic health records (EHR) information. This paper investigates the key challenges, techniques used, technologies used, privacy, security algorithms and future directions of Big Data Analytics in the healthcare and medical data.

Index Terms: Healthcare Management, Cloud Computing, Data Storage, Big Data Analytics, Privacy, Security, Medical

1 INTRODUCTION

For the global healthcare industry, changes will be occur every day. In reality, the digitization of healthcare and patient's information is basic or fundamental and dramatic change in the business, clinical and operational models and the world of the economy in general for the future. Cloud Computing for Big Data delivers promises, benefits and offers possible that is excellent medical data, yet it makes numerous hurdles and difficulties. Indeed, concerns regarding privacy and security of big healthcare data are increasing year-by-year. Big data analytics application offers extensive understanding that can be discovered from the enormous quantity of data available. The data provided through Big Data Analytics methods should provide patients, clinicians, and health policy advisors with extensive advantages. Cloud Computing in healthcare will make it much easier for a patient to collect all the distinct healthcare data while moves the patient from one hospital to another hospital, making it simply to manage and track patient data. Security, Flexibility, parallel processing, virtualization and scalability of resources are its benefits. Cloud computing can reduce automation expenses through automation and maintenance of infrastructure, while improving operational efficiency and user access.

2 LITERATURE REVIEW

The increase in data every day, the amount of devices those are related to internet enabled for data generation and the dropping data storage costs themselves, which make the data accessible for nearly no cost to everyone. Healthcare or medical is the most primary area in the world. Big data analytics has making creative presentation compared to standard means. Since nurses are unable to take their medical document with them at any moment, confidential patient data must be held safe. Since patients are unable to carry their medical document with them every time, confidential patient data must be kept secure. The data could be stored and encrypted in cloud environment for this purpose in order to protect sensitive data privacy. But support for activities such as cloud search for keywords should be given.

1. In Senthil Kumar et al. (2018) [1], privacy preservation of heathcare data is done where the sensitive data is stored and encrypted in the cloud environment. The encrypted data could be stored in cloud environment using the walrus and attribute based encryption (ABE) is performed.

2. E. Shanmugapriya et al. (2019) [2], Proposed method investigates privacy and security with hybrid cloud computing. It was used bilinear pairing protocol and also used authenticated key management system. The Proposed method will provide the less computation cost, time consumption, and computational complexity compared to existing method.

3. Shin et al. (2014) [3], reviewed multiple healthcare application security models and tried to see how data leakages could be protected. To guarantee the security and privacy in medical data or healthcare, assessed security that is multiple. It investigates Role Based Access Control (RBAC) security mechanism [24] to discover solutions to recognized safety problems in electronic health. They developed an integration platform for the u-healthcare service where an expanded RBAC model was deployed. For any distributed environment, the model is not appropriate. As a consequence, there are restricted applications in the solution given. The application also does not consider the amount of customers expanding.

4. Gajanayake et al. (2016) [4], developed a unique privacy and guaranteed e-Health access control architecture. It combines the three type of security models. Those are Role Based Access Control (RBAC), Mandatory access Control

- *Mr.Sreehari Kundella is currently pursuing Ph.D. Degree in Computer Science at Vels Institute of Science, Technology & Advanced Studies (VISTAS), and Chennai, India. PH-91-9642452410.*
  E-mail: Kundella.sreehari@gmail.com
- *Dr.R.Gobinath is currently working as an Associate Professor at Vels Institute of Science, Technology & Advanced Studies (VISTAS), and Chennai, India. E-mail: iamgobinathmca@gmail.com*

(MAC) and Discretionary Access Control (DAC) [10] to develop a new architecture that enables patients and hospitals to determine and setup the rights to accessing the data. The main disadvantage of the structure is employed for standalone protection model to achieve health record electronic demands.

5. Rezaeibagha and Mu (2016), In Electronic Health Record (EHR) created a new access-control structure to tackle safety and privacy challenges. The framework is employed hybrid clouds plus the transformation of access control policies to make sure reliable and access that is reliable and authorization for sharing of the data between the various healthcare providers. Some of the cryptographic blocks were implemented using of access control policy conversion to address the multiple EHR users with distinct access privileges and authorization in distinct cloud settings in order to make the model effective. The primary drawback of this structure is its inability to provide space for user development. It does not offer scalability space because there is a restricted amount of users.

6. Wang et al. (2017) [6], improve the attribute-based encryption (ABE) introduced. The enhanced KP ABE (Key Policy ABE) model and CP ABE (Cipher Text Policy ABE) based auxiliary input schemes are provided. The enhanced model also takes into account the encrypted leakage (leakage of randomness) in front of other auxiliary input models while conducting the comparison. In addition to the modified Goldreich - Levin theory provides an enhanced powerful extractor.

7. Qinlong Huang et al. (2017) [7], It investigates on the health care that is secure the data for smart towns and cities. It really is in line with the ABE and IBBE. The framework enables the data proprietor to permit healthcare analyzers to gain access to information by re-encoding both ABE-protected data which are medical IBBE safeguarded social data to IBE protected information that offers a remedy for cooperation between distinct service companies. To diminish the overhead computing of resource constrained mobile devices, our bodies adopts encryption that is outsourced decryption building, which can delegate all the cost of computing up to a cloud serve.

8. Jong Wook Kim et al.(2018)[8], This paper explores a privacy that is brand new for collecting private health information streams that is understood to be short-term data collected at pair of periods by using local differential privacy (LDP)[8]. A tiny amount of salient information acquired from the complete health data stream in particular, data factor utilizes a specified LDP privacy spending plan because of the recommended technique to be accountable to a data collector. The proposed strategy first describes a quantity that is small of points from a data factor's whole health information stream, disrupts these recognized data points under LDP, then reports[8] to a data collector the disrupted data in place of reporting all of the data within the flow.

9. Mehedi Masud et al. (2017) [13], this paper presents a guaranteed in full regarding the fly exchange protocol for the cloud environment that is sharing Pairing Based Cryptography (PKI). Every cloud computes a key session that should be secret, for trading private information by computing a mix for each data session that is sharing. The secret key is produced dynamically utilizing arbitrarily produced parameters for each and every new data session.

3 SECURITY IN BIG DATA ANALYTICS WITH RESPECT TO CLOUD COMPUTING

Cloud computing and big data analytics will be a critical phases from development. Data privacy and security both are essential problems in the cloud due to the open environment with extremely controls individually that is bound. The information is stored in a spot that is key while the cloud storage host, where the information is processed on the cloud servers, this provides you with your client rely upon the service provider and data safety. The solution level agreement needs to be standardized to achieve trust between companies and customer Cloud customer data security differs with protective demands/requirements. This paper investigates in the existing security issues. Those are described below in table format.

TABLE 1
EXISTING SECURITY ALGORITHMS IN BIG DATA AND CLOUD COMPUTING

Author	Year	Title	Security Algorithm	Description
Senthil Kumar R,Latha Parthiban	2018	Privacy Preservation In Big Data With Encrypted Cloud Data Storage Using Walrus	Attribute Based Encryption (ABE)	Walrus, It is a storage service in Eucalyptus for storing the data. It will store the data in cloud in the form of buckets. The encrypted information will be stored in cloud using the walrus and by performing the ABE.
R. Kavitha, E. Shanmugapriya [2]	2019	Medical big data analysis: preserving security and privacy with hybrid cloud technology [2]	Bilinear pairing cryptography	Proposed method investigates privacy and security with hybrid cloud computing. It was implemented by bilinear pairing protocol to analyze the big data and using authenticated key management system. The Proposed method will provide the less computation cost, time consumption, and computational complexity compared to existing method.
Licheng Wang ,Qinlong Huang, and Yixian Yang [7]	2017	Secure and Privacy-Preserving Data Sharing and Collaboration in Mobile Healthcare Social Networks of Smart Cities [7]	ABE, IBBE [25]	It focuses on the healthcare data and social sharing data and collaboration smart cities. It is developed based on ABE and IBBE [25].
Dr. K. Baskaran	2016	Cryptographically	Revocable Multi	Proposed method will be used Revocable Multi

and Dr. Ragesh G. K. [20]		Enforced Data Access Control in Personal Health Record Systems[20]	Authority Attribute Set Based Encryption (R- MA-ASBE)[20]	Authority Attribute Set Based Encryption (R-MA-ASBE) [20]. Each patient data would be encrypted, prior to uploading the data in the cloud server.
Tejaswini L and Dr. Nagesh H.R. [14]	2017	Study on Encryption methods to secure the Privacy of the data and Computation on Encrypted data present at cloud [14]	Mohomorphic encryption	It uses the homomorphic encryption. The data stored at cloud can kept private and also the computation on the cipher text can be achieved. Since the healthcare record needs to be kept secure and private this approach can be used.
Jin Sun, Xiaojing Wang, Shangping Wang, Lili Ren[21]	2018	A searchable personal health records framework with fine-grained access control in cloud-fog computing [21]	Search Encryption (SE) Technology and Attribute-Based Encryption (ABE)	The proposed article combines the ABE and SE for implementation of keyword search Function and ability to access the control that is fine-grained. If the trapdoor and keyword match both are successful, then a cloud host provider gives the results based on the search to the [21] individual based the search requirements
Ling Liu, Rui Zhang and Rui Xue [23]	2017	Searchable Encryption for Healthcare Clouds: A Survey [23]	Searchable Encryption	This paper defines the encryption that is searchable encryption the healthcare applications.
Tingting Zhang, Yang Ming [22]	2018	Efficient Privacy-Preserving Access Control Scheme in Electronic Health Records System [22]	Privacy Preserving Access Control (PPAC) Mechanism	This Paper defines new method called PPAC Mechanism [26] for Medical Records. It utilizes the attribute based signcryption method to signcrypt the data.

TABLE 2
MEASUREMENTS USED IN BIG DATA IN CLOUD

Author	Year	Title	Algorithm	Parameters/ Measurements Used	Description
R. Kavitha, E. Shanmugapriya [2]	2019	Medical big data analysis: preserving security and privacy with hybrid cloud technology	Bilinear pairing cryptography	Quality, Cost, Time	Proposed method investigates privacy and security with hybrid cloud computing. It was using bilinear pairing protocol to analyze the big data and using authenticated key management system. The method will provide the less computation cost, time consumption, and computational complexity compared to existing method.
Licheng Wang, Qinlong Huang, and Yixian Yang[7]	2017	Secure and Privacy-Preserving Data Sharing and Collaboration in Mobile Healthcare Social Networks of Smart Cities[7]	ABE, IBBE	Size, user's Identity, Public Key, Message	It focuses on the healthcare and social sharing data in mobile healthcare for smart cities. It defines using the ABE and IBBE [7].
Jong Wook KimID, Beakcheol Jang, Hoon Yoo	2018	Privacy-preserving aggregation of personal health data streams	Local differential privacy	Data, Rate, Time, Total Number of users	The proposed methods defines a minimal amount of data points from a data contributor's entire data streams which are related to health, disrupts these recognized data points under LDP and reports to a data collector the disturbed data in the place of reporting the complete data into data streams.[8]
Mehedi Masud, M. [13]Shamim Hossain	2017	Secure data-exchange protocol in a cloud-based collaborative health care environment	High Level Protocol Specification Language (HLPSL), pairing-based cryptography	System parameter, Session parameter, Secret Shared Parameter, Shared Secret Identity, Message	This paper introduces a guaranteed on-the-fly and secured data-exchange protocol in the cloud sharing data environment using pairing based cryptography (PKI).
Yiping Wen, Jianxun Liu, Wanchun Dou, Xiaolong Xu, Buqing Cao, Jinjun Chen [15]	2017	Scheduling workflows with privacy protection constraints for big data applications on cloud	Multi-Objective Privacy-Aware workflow scheduling algorithm(MOPA) [15]	computational cost, data transfer cost and storage cost, Time, Data,	This model, the issue of scheduling workflows with restrictions on privacy protection while minimizing both execution time and financial cost. A new programming algorithm called MOPA is suggested for Multi-Objective Privacy-Aware. We contrasted it with two other algorithms as well. The experimental finding indicates that the MOPA algorithm can come up with better alternatives than others

Kingsford KISSI MIREKU, Fengli ZHANG, Komlan GBONGLI, Isaac Edem DJIMESAH	2018	Privacy Preservation Data Publishing of Health Insurance Scheme in Ghana using Identity Base Encryption Scheme	Identity Base Encryption	Membership, Identity, data	This paper investigates, when submitting insurance claims, used the IBE scheme to delink the SM identity from the healthcare data. In the process of releasing data on big data storage by the National Health Insurance Authority (NHIA), it is unsafe for these scientists, the press agencies or the Ghana statisticians (these are their primary users of data) to have access to the identities of the members of the system for their respectively.
A.M. Vengadapurvaja, G. Nisha, R. Aarthy, N. Sasikaladevi	2017	An Efficient Homomorphic Medical Image Encryption Algorithm For Cloud Storage Security	Homomorphic Encryption	Patient's Data like clinical records, scanning and x-ray images, public and private keys	This article proposes an effective homomorp hic encryption algorithm for encrypting medic al images and carrying out helpful activities without violating the confidentiality
Yingjie Xia, Xuejiao Liu, Wei Yang, Fengli Yang[19]	2016	Secure and Effi-cient Querying over Personal Health Records in Cloud Computing	Attribute Based Encryption (ABE), CP-ABE approach	Security Parameter, User Data, Key	It is used ABE algorithm in healthcare about Personal Health Records (PHR) sharing in cloud.

4 DIFFERENCE BETWEEN THE SECURITY AND THE PRIVACY

Big data analytics security and privacy both are major problems. Privacy is described as capable of protecting the data in the area of healthcare data that is recognizable personally. It is focused on the usage and managing the data about individuals, making policies and setting the requirements for authorization to ensure the personal data of patients are collected, shared and used in the proper and right way. While security is described as protection from users unauthorized access, which are include explicit reference to accessibility and integrity. It focuses on defending data from harmful assaults and for profit stealing the data. While security is an essential to data protection, it is not enough thing to address privacy of the data.

TABLE 3
DIFFERENCE BETWEEN THE SECURITY AND THE PRIVACY

Security	Privacy
It is a concern about the data's confidentiality and more over integrity and accessibility.	It is a proper usage of the data about the user's perspective.
Different techniques will be used such as Encryption, Firewall, etc. to prevent the data compromising from an organization's network of technology or vulnerabilities.	The organization cannot sell the data to the third parties without patient / user approval.
Ensure confidentiality or safeguard a company or organization	It concerns the right of the patient to protect the data from any other third party
Security provides the capacity to be confident of compliance with choices.	Privacy is the capacity to determine what an individual's data is going to and where.

5 CONCLUSION

This paper investigates, the privacy and protection/security algorithms in huge amounts of the data and cloud processing in health by speaking about formulas which can be present practices found in which health businesses are usually extremely useful in complete protection and privacy. In this part, dedicated to techniques and strategies provided in numerous documents on the focus and limits also offered protection and privacy issues in big data and cloud processing along with present technologies within the framework of huge data in healthcare privacy and protection. In this paper, primarily evaluated the strategies of privacy conservation recently found in health care and resolved how practices of anonymization and encryption were used for data protection in health and described their constraints. Big Data analytics that is uncover valuable. In the years which can be impending cloud computing will continue to produce and accelerate within the health care business. In this framework, as our future direction, prospects contains attaining privacy that is efficient safety choices into the age huge all about health. Types of privacy should also be enhanced. Eventually, Cloud Computing's Big Data Analytics area is an important tool to rise the researchers to deliver alternatives which can be sufficient the ever-emerging issues into the sector.

REFERENCES

[1] [Senthil Kumar R, Latha Parthiban. PRIVACY PRESERVATION IN BIG DATA WITH ENCRYPTED CLOUD DATA STORAGE USING WALRUS, International Journal of Pure and Applied Mathematics, Volume 119 No. 15 2018, 1833-1842.

[2] E. Shanmugapriya, R. Kavitha. Medical big data analysis: preserving security and privacy with hybrid cloud technology, Springer-Verlag GmbH Germany, part of Springer Nature 2019, s00500-019-03857-z.

[3] Shin M, Jeon H, Ju Y, Lee B, and Jeong S. Constructing RBAC based security model in u-healthcare service platform. Sci World J 2014; 1–13.

[4] Gajanayake R, Iannella R, Sahama T. Privacy oriented access control for electronic health records. E-J Health Inf 2014; 8(2):175–86.

[5] Rezaeibagha F, Mu Y. Distributed clinical data sharing via dynamic access control policy transformation. Int J Med Inf 2016:25–31.

[6] Z. Wang, C. Cao, N. Yang, and V. Chang, "ABE with improved auxiliary input for big data security," Journal of Computer and System Sciences, vol. 89, pp. 41–50, 2017.

[7] Qinlong Huang, Licheng Wang, and Yixian Yang. Secure and Privacy-Preserving Data Sharing and Collaboration in Mobile

Healthcare Social Networks of Smart Cities, Hindawi Security and Communication Networks Volume 2017, Article ID 6426495, 12 pages.
[8] Jong Wook Kim, Beakcheol Jang, Hoon Yoo. Privacy-preserving aggregation of personal health data streams, 2018.
[9] Nureni Ayofe Azeez, Charles Van der Vyver. Security and privacy issues in e-health cloud-based system: A comprehensive content analysis, Egyptian Informatics Journal, 2018
[10] Karim Abouelmehdi, Abderrahim Beni‐Hessane and Hayat Khaloufi. Big healthcare data: preserving security and privacy, Journal of Big Data (2018) 5:1.
[11] Sudipta Chandra, Soumya Ray, R.T.Goswami. Big Data Security in Healthcare: Survey on Frameworks and Algorithms, 2017 IEEE 7th International Advance Computing Conference (IACC), PP. 89 – 94.
[12] Priyank Jain, Manasi Gyanchandani and Nilay Khare. Differential privacy: its technological prescriptive using big data, Journal of Big Data, 2018 5:15.
[13] Mehedi Masud.M. Shamim Hossain. Secure data-exchange protocol in a cloud-based Collaborative health care environment, Springer Science+Business Media, LLC 2017, s11042-017-5294-5.
[14] Dr.Nagesh H.R., Thejaswini L. Study on Encryption methods to secure the Privacy of the data and Computation on Encrypted data present at Cloud, IEEE, 2017, 978-1-5090-6399-4/17, pp. 383-386.
[15] Yiping Wen, Jianxun Liu, Wanchun Dou, Xiaolong Xu, Buqing Cao, Jinjun Chen. Scheduling workflows with privacy protection constraints for big data Applications on cloud, Future Generation Computer Systems, 0167-739X/© 2018 Published by Elsevier B.V.
[16] L. Wang, R. Ranjan, J. Kołodziej, A. Zomaya, L. Alem, Software Tools and Techniques for Big Data Computing in Healthcare Clouds, Futur. Gener. Comput. Syst. 43 (2015) 38-39. DOI: 10.1016/j. future.2014.11.001.
[17] Kingsford KISSI MIREKU, Fengli ZHANG and Komlan GBONGLI, Isaac Edem DJIMESAH, Privacy Preservation Data Publishing of Health Insurance Scheme in Ghana using Identity Base Encryption Scheme, Proceedings of the 9th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2018), pp. 113 – 118.
[18] A.M. Vengadapurvaja, G. Nisha, R. Aarthy, N. Sasikaladevi. An Efficient Homomorphic Medical Image Encryption Algorithm for Cloud Storage Security, Procedia Computer Science 115 (2017) 643–650.
[19] Xuejiao Liu, Yingjie Xia, Wei Yang, Fengli Yang, Secure and Efficient Querying over Personal Health Records in Cloud Computing, Neurocomputing (2017).
[20] Dr. Ragesh G. K., Dr. K. Baskaran, Cryptographically Enforced Data Access Control in Personal Health Record Systems, Procedia Technology 25 (2016) 473 – 480.
[21] Jin Sun, Xiaojing Wang, Shangping Wang, Lili Ren, A searchable personal health records framework with fine-grained access control in cloud-fog computing, PLOS ONE (2018).
[22] Yang Ming and Tingting Zhang, Efficient Privacy-Preserving Access Control Scheme in Electronic Health Records System, Sensors 2018, 3520.
[23] Rui Zhang, Rui Xue, and Ling Liu, Searchable Encryption for Healthcare Clouds: A Survey, 1939-1374 (c) 2017 IEEE.
[24] Nureni Ayofe Azeez, Charles Van der Vyver."Security and privacy issues in e-health cloudbased system: A comprehensive content analysis", Egyptian Informatics Journal, 2019
[25] Isma Masood, Yongli Wang, Ali Daud, Naif Radi Aljohani, Hassan Dawood. "Towards Smart Healthcare: Patient Data Privacy and Security in Sensor-Cloud Infrastructure", Wireless Communications and Mobile Computing, 2018
[26] M.A.P. Chamikara, P. Bertok, D. Liu, S. Camtepe, I. Khalil. "An efficient and scalable privacy preserving algorithm for big data and data streams", Computers & Security, 2019