An Efficient Privacy Preserving Approach for e-Health

Supriya Menon M[1]
Research Scholar
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddesvaram,AP, India

Dr. Rajarajeswari Pothuraju[2]
Professor
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddesvaram,AP, India

***Abstract*—Immense Procreation of large amounts of data in medical field and health care domain, benefitting society is at risk with sensitive attributes being disclosed. Access to Medical Information made feasible over internet with an intension of serving the people related to medical community is triggering a challenge for researchers in norms of Privacy and security. The medical data at cloud is vulnerable to unpredictable threats with evolving technology, and the threat landscape sounds resilient with sensitive attributes. In this contemporary stretch, Organizations fail to hold the reputation and are unable to preserve public confidence. The austerity of sophisticated security attacks compromise the privacy of patient data and security of healthcare units. The fruitful approaches by several researches and practitioners provided an up heal resolutions, but the demand for an optimal solution remains unanswered. In this paper we present a solution for addressing the security issues in health care management. We propose a hybrid framework using enhanced Attribute Based Encryption (ABE) with Anonymity approach based on access primitives of sensitive attributes. The proposed mechanism is evaluated in terms of performance, encryption time, decryption time and Memory utilization using Jsim simulator which envisage drastic performance expedition in the presented model.**

***Keywords*—e-Health; Attribute Based Encryption (ABE); secure hash algorithm (SHA-1); anonymity; privacy; sensitive parameters**

I. INTRODUCTION

Medical domain owes to be attractive region for researchers, challenging them in various aspects like disease prediction, Drug prediction, Drug Repositioning and many more. The recent research focused on disease and treatment prediction using medical repositories accessed and published over distributed environments [1]. This Medical records accessed by several authorities put forth's queries questioning the security and privacy of health care data. The cause entailing such need divulged from the fact that health care data is outsourced for various reasons at the risk of compromising privacy requirements like confidentiality [2, 3], Integrity, keyword privacy, authentication audit ability and further. Additionally evolving technologies of information and communication attracts medical domain for integrating health data with technology [4] from domain's like Hospitals, health insurance firms and research laboratories leading to e-Health.

E- Health, an attractive domain in recent times that overlaps public health and medical informatics with corporate sectors over internet aims at improving data analysis of health care data locally and worldwide. The cloud successfully offers few advantages over network primarily in enhancing patient care by supporting interaction [5] with healthcare authorities and availability of patient data for analysis and diagnosis [6]. e-Health also offers support for medical research in disease treatment prediction with extended monitoring of epidemics. Further, it helps in cost reduction for engaging expensive hardware, software and data storage at premises.

Eventually e-Health showcases few pullbacks like complexity in interoperability i.e. lacking standard for synchronization, security and privacy issues [20] in shared and public environments [8], regulation controversies related to social and valid frameworks and reliability considerations. They even need to work hands with sensor networks [9] involving data collection. The Healthcare providers in practical are surrounded with several risks [10] from digital technology on cloud despite they encircled advantages.

Among aforementioned issues security and privacy challenges demand utmost attention for realization of its effective utilization. Data confidentiality, authentication and Integrity are at risk in distributed environment. The goal of medical data shared and stored over internet is to provide consistency and high level of security. Despite numerous cryptographic and non-cryptographic methodologies available for enduring security and privacy of e-health data [11, 12], few unturned grains are hindering hurdles constraining performance. Our approach proposes a way forward for contending the security and privacy gaps in e-health [13].

The proposed architectural framework attains the goal of security using a Hybrid ABE as well as provides selective access to records based on user predefined access policies like authorized uses, restricted users and un authorized users. The Hybrid ABE provides efficient performance using secure hash algorithm (SHA-1) and Anonymity approaches. The hybrid approach promises high degree of availability, reliability, and efficiency in protecting patient sensitive information [14] upon implementation. Explicitly the architecture gives room for desired authentication to medical archives with extended control for medical stakeholders in general and emergency scenarios on demand.

II. Related Work

Puneeth Saran et al 2020 contributed to a qualitative research regarding role of cloud computing in medical field and proposed a method to increase the security of medical records in cloud. Gaozhiqiang et al 2015,proposed a cloud based remote health care consisting of portable medical devices, intelligent terminals ,cloud platforms where user can access their health data via internet. Inderpreet Singh et al 2019, presented a model for grouping adaptable e-health care services depending on distributed computing environment which showcases high correctness rate for secure information access. KnutHaufe et al 2014 presented a framework for security of health care records stored in cloud and identified ISMS process that needs to be focused for future research. R.Anitha & Saswathi Mukherjee 2014 proposed a novel framework -generating cipher key from attributes of metadata created by DCMI standard using patients medical record. Luliana Chiuchisan et al 2017 made a detailed survey of security measures involved in health care management and proposed health care system that monitors rehabilitation of patients with Parkinson's disease. Alexandru Soceanu et al 2015 presented encryption scheme and attribute based framework with encryption process relying on ARCANA tool for secure hierarchical access. Yaza-Al-Issa et al. 2019 reviewed with regard to cloud computing services in health care management and privacy concerns for health care providers and reiterated that only few concerns of security are addressed. Shekha Chenthara et al. 2019 contributed to intensive survey about HER (Electronic Health Record) security and privacy, EHR cryptographic and non cryptographic approaches in IEEE, Science direct, Google scholar, Pub Med and ACM library. Nureni Ayofe Azeez & Charles Van Der Vyer, 2018reviewed 110 original articles, figured out various models adopted with their standard definitions on e-health and proposed secured architecture for e-health to provide privacy between health care providers and patients. Isma massod, 2018, proposed six step generic framework for patient physiological parameters, privacy and security in sensor supported cloud infrastructure with performance evolution in research. Ronald Glasberg et al., 2014, analyzed risks and crisis for health care providers in holistic way, taking organizational and human aspects into account. Shyh -Wei Chen et. al., 2016, proposed architecture of patient centered personnel health record to manage patient health information and health reports with cloud based secure transmission. Alexmu-Hsing Kuo et. al., discussed in detail about health care and considers four aspects to analyze the challenge of cloud computing model. Ramzi. A et.al. presented recovery algorithm using concept of matrix in health care management and evaluated the performance against various techniques. Panjunsun, 2019, proposed privacy protection framework by reviewing challenges and solutions of data security in detail. Uma narayanan et al. 2020, proposed novel system architecture called security authentication and data sharing in cloud (SADS –cloud) including SHA-3 hashing algorithm for registered data owners. Ijaz Ahmad Awan, 2020 proposed framework deploys AES with 16, 32, 64, and 128 plaintext bytes enhancing security and minimizing resource utilization in computational clouds. Arafat Al-Dhaqm et al., 2020, gave a detailed review on DBFI – Database Forensic Investigation and proposed harmonized DBFI process using systematic approach with higher certainty. SupriyaMenon M and Rajarajeswari P, 2018, reviewed privacy issues of personalized and context aware privacy and proposed a model for context aware privacy. Jitendra Kumar and Ashutosh Kumar Singh, 2017, came up with a workload prediction model using Long short term memory (LSTM) and tested over three web log datasets proving enhanced accuracy by proposed approach. Supriyamenon M and Rajeswari P, 2020, addressed the complications related to drug repositioning and came up with a hybrid ACO approach enhancing Drug consumption similarities for better repositioning addressing the need for secure patient data. Ma, H., Zhang, R., & Yuan, W, 2016, contributed a mode;l for ABE based Anonymity for Identity revelation.

III. Security Preliminaries

A. ABE

ABE is an encryption scheme, where the generated cipher text is an outcome relying on user private key and attributes of user data. This public key encryption technique renders plaintext at requested site with decryption supported upon attribute matches of user key and cipher text attributes [15] from attributes of metadata. Although initially introduced in its basic form, exploring amendments of attribute based encryption [16] with multiple authorities involving in user private key generation are also available. ABE has its wide spread usage in several areas like vector driven search engine interfaces, log encryption avoiding log encryption with all recipient keys.

There are two forms of ABE one for key policy KP-ABE and other for cipher policy CP-ABE. The KP-ABE generates user private depending on access tree related to user privileges and encrypting over a set of attributes using algorithms like AES [17]. However cipher text based ABE encrypts user data and attribute with secret keys generated from access trees.

ABE rising to be a well preferred mechanism is surrounded with overwhelming challenges like in efficient attribute revocation mechanism, improper key co-ordination, key escrow deficiency and issues related to key revocation mainly for healthcare systems [18] .Few extended problems in the path of ABE is its centralized concept. The need for a centralized body or authority participating in private key generation, makes ABE encounter the flaws due to lack of decentralization. These risks bring down the performance of ABE. One more factor of concern affecting the ABE is speed, which downtrends compared to others due to delay of policy tree construction and computational delay at decryption site also adds upon the issue.

B. SHA-1

The secure hash algorithms enable the determination of Message Integrity that facilitate creation and validation of digital signatures. Digital signatures provide secure security service of Authentication [19] hereby avoiding Denial attacks and repudiations both at source and destination. SHA-1 belongs to the family of secure hash algorithms that generate a hash value known as message digest to facilitate security [21]. It promises its wide spread excellence in several security

protocols, mail protocols, TTL, SSL, IPSec and many more. The basic version of the algorithm produces a 160 bit message digest which well prevails against Brute force attack. This variant is considered to be the fastest one but more prone to collision problem, those were overruled in the successor variants. Few well known variants of secure hash algorithms are SHA -2 and SHA-3. The former uses a set of 6 hash functions with digests of size 224, 256, 384, and 512 bits. Among the aforementioned digests SHA-256 and SHA 512 exhibit uniqueness in the sense of computing with 32 bit and 64 bits respectively. They project the variation in the basic shift and additive operations performed. SHA-2 being advanced faced a strong battle to take over its Predecessor. The later addressed SHA-3 by NIST provides compatibility with the former.

C. Anonymization

Anonymization is a process that aims at encapsulating identifying information in a way intending privacy protection. Hence the original data remains anonymous enabling data sharing and transmission among agencies reducing risk of unwanted disclosures [22]. Despite such secure transformations anonymous data never promise to anonymous over time. Several approaches and clever techniques exist that disclose data leading to be de-anonymizated. To handle all such loop holes, several forms of Anonymity are available like k-Anonymity, *I*-Anonymity, t- closeness, *p*-sensitivity and many more variations. In k-anonymity, anonymization is a key feature using certain cryptographic hashing. K-Anonymity further has its extension to an (α, k)-anonymity model for privacy preserving data publishing, where α being a fraction and k an integer. The frequency of sensitive value is no more than α. It aims at data security and privacy with further extension to Human and Societal aspects of security and privacy.

IV. Proposed Approach

Huge amounts of data filling the health care repository is triggering several challenges in due response to providing services. These services claim that cloud computing techniques provide everything as a service i.e. storage as well as security as a service. The major issue of concern is medical confidentiality, portraying the healthy relationship of trust among patients and doctors. The medical data stored in cloud is at high risk of being vulnerable to attacks with irretrievable loss to users with their sensitive data dumped at entrusted servers. With an intension of addressing the above mentioned issues related to data privacy we propose an hybrid approach that resolves the complications in data transmission and provide security.

Phases in proposed approach are discussed below.

Phase 1:

This phase of the proposed system initiates with generation of metadata for the patient records. The attributes in patient records are analyzed and access control structure is defined considering different threshold parameters for various groups of users using ABE approach. Certified attributes defined in the access policy determines which block of plain text should be decrypted for the users with predefined threshold credentials. Elicited from the defined access policies, users are assigned access permissions to the available records.

Phase 2:

The medical records blocks are encrypted considering four randomized algorithms in ABE as Setup, Key generation, Encryption and Decryption.

Setup: At initialization the system generates 2 groups GR1 and GR2 based on security parameters with p prime value, t threshold and b bilinear pairs.

The centralized authority generates master key M_K and public key P_K by randomly selecting x, u1, u2,…u_n Z_q where q is the prime number and Z_q multiplicative modulo.

Key generation: The authorized authority generates secret key S_k for users by using SHA1 with modified feistel structure where SHA1 converts attributes into matrices considering m rows and n columns, where m is the number of attributes and n is the size of SHA output.

The algorithm for key generation is presented below.

1. Initiates by reading sensitive attributes and inputting them to SHA-512.
2. Resultant Matrix A_{mxn} is further divided into A_1, A2, A3, and A_4.
3. Produce Le_{mxn} and Ri_{mxn} by combining A_1, A_3 and A_2, A_4 respectively.
4. Left and Right values of Feistel network undergo following computations.
   - Bifurcating Ri_{mxn} to equal partition matrices R_{ia} and R_{ib}.
   - Apply transposition resulting in R_{ia}nxm and R_{ib}nxm and add them to Q_{mxn} .
   - RO_{mxn} a resulting transpose of Q_{mxn} .
   - Revised Le_{mxn} is RO_{mxn} and Ri_{mxn} is previous value of Le_{mxn} until n holds old value.
   - Output $K_{mxn} = Le_{mxn} \| Ri_{mxn}$.
5. Process terminates.

Encryption: Sender encrypts the message with key extracted from attributes.

C_i = Encryption (P_k, PT, A) where P_k public key from attributes A for plaintext PT.

Decryption: The Receiver decrypts Ciper $C_{i\ using}$ the Secret key S_k generated by SHA1.

ABE explicitly supports threshold operations on attributes to specify permitted access control structures to the users of different groups.

Phase 3:

Among different groups of users with certain attribute combinations, the limited access groups of users considered as restricted users are subjected to feasible Anonymity technique

with lower distortion. The k- anonymity technique preferred avoids identity disclosure.

Algorithm for (a, k)Anonymity :

Input: Raw table

Output: Hybrid Anonymity table.

1. Initialization stage

Generate user for input vector and for array of users considered 1,2,3,…n, compute dissimilarity matrix DMT by calculating distances.

2. Anonymization at Client side

1. Compute DMT.
2. Assign false to all points and select a point p with C_c as centroid of user c_i and mark it as true.
3. Consider false points as minimum distance from C,with social attributes S_A.
4. Add above considered point to P, and check the frequency with Anonymization parameter k.

If frequency of (S_A) < k
{
Consider and adjust centroid;
}
Else
Abort;

4. Repeat until all points in c are verified and return.
5. Group unassigned points to nearest user and ensure user satisfying (1,K1) anonymity.

3. Anonymization at Server side

1. Consider the nearest user pairs p1 and p2 in client side matrix.
2. Combine p1and p2.
3. Size of u=p1+p2.
4. Compute representative vector R* u using tree access structure T.
5. T_y denotes sub-tree where y is root in tree T.
   When root node is r in T= T_r Attribute set =e that confirms to access policy of T_y
   Then T_y (e) =1
   If y is a leaf node and attribute attr (y)
   T_y (e) =1
   Else
   Validate sub- nodes.
6. Repeat process until each user satisfies (a2, k2) anonymity.

Based on the Anonymity levels attained from proposed technique predefined threshold attributes in access tree structure are sent in Plain text and other blocks are anonym zed.

V. Performance Evaluation and Analysis

Our Proposed Hybrid ABE Approach projects efficient performance with respect to time and Memory Utilization when compared with existing techniques like Common Database Forensic Investigation Process (CDFIP), Real-time Operational Data Base (RODB) Extraction–Transformation–Loading (ETL), and Long Short Term Memory (LSTM) [7] to Recurrent Neural Network (RNN). The Simulation parameters considered for Implementation of the proposed approach are shown in the Table I.

To evaluate the performance of Hybrid ABE, Encryption time, Decryption time and Memory Utilization are considered.

Encryption time: This computes the throughput of the encryption scheme with respect to user instances and encryption time.

Decryption time: This computes the throughput of the decryption scheme with respect to user instances and decryption time.

Memory Utilization: This evaluation parameter projects the average utilization of system memory in bytes for different user instances.

Table II shows an improved performance of time for different user instances using Hybrid Approach against existing methods opted.

Fig. 1 depicts improved performance of the Proposed Approach contradicting existing Approaches mentioned.

TABLE I. Simulation Parameters

PARAMETERS	VALUES
Simulator	Jsim
Simulator Time	120 s
Proposed Protocol	Hybrid Approach
Number of user instances	10 - 500

TABLE II. Performance Comparison

No. of user instances	CDFIP	RODB & ETL	LSTM to RNN	Hybrid Approach
10	4.3	3.7	3.6	1.9
30	5.4	4.8	5.2	3.4
50	6.4	5.4	4.6	3.7
70	7.3	6.2	7.1	4.6
100	8.6	7.4	6.3	7.3

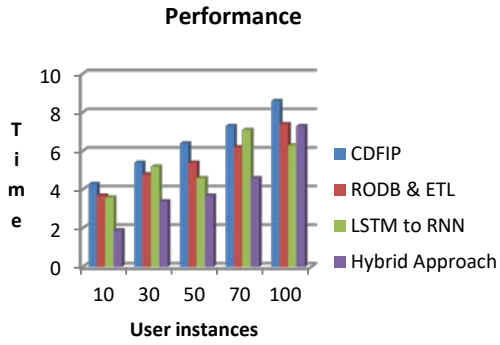

Fig. 1. Performance Graph.

Table III displays a comparison of encryption time for different user instances against Hybrid and existing approaches.

TABLE III. ENCRYPTION TIME COMPARISON

No. of user instances	CDFIP	RODB & ETL	LSTM to RNN	Hybrid Approach
100	4.3	3.7	3.6	3.5
200	5.4	4.8	5.2	4.1
300	6.4	6.7	5.2	4.3
400	7.3	6.2	7.1	4.6
500	8.6	7.4	6.3	7.3

Fig. 2 shows a clear comparison of the reduced Encryption time for proposed in comparison to the existing approaches as user instances keep varying.

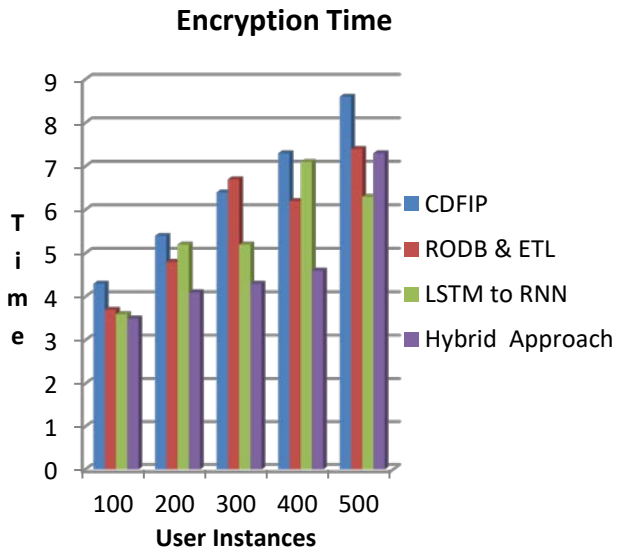

Fig. 2. Encryption Time.

Table IV shows a comparison of decryption time for various user instances against Hybrid and existing approaches.

TABLE IV. DECRYPTION TIME COMPARISON

No of user instances	CDFIP	RODB & ETL	LSTM to RNN	Hybrid Approach
100	3.7	4.7	4.2	3.8
200	4.2	5.6	4.8	3.6
300	3.7	7.4	5.3	4.3
400	6.3	5.82	4.6	4.7
500	5.7	6.4	6.8	5.3

Fig. 3 depicts leveraged decryption throughput of Hybrid approach compared to other approaches.

Table V depicts the memory utilization of different approaches against proposed approach taking into consideration the varying instances of users.

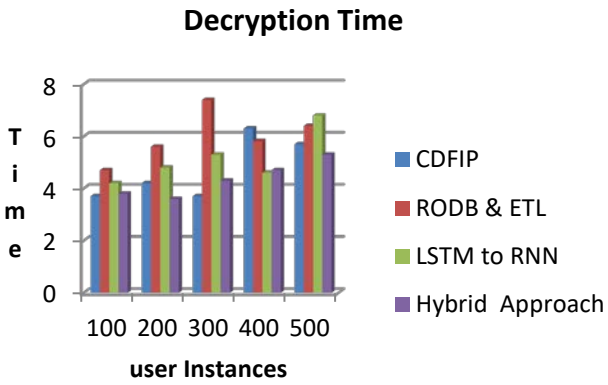

Fig. 3. Decryption Time.

TABLE V. MEMORY UTILIZATION

No of user instances	CDFIP	RODB & ETL	LSTM to RNN	Proposed Approach
100	3541	3642	3876	2759
200	4216	4326	4216	3124
300	6245	5974	3654	3926
400	11021	6785	5243	4146
500	23542	7853	5674	5214

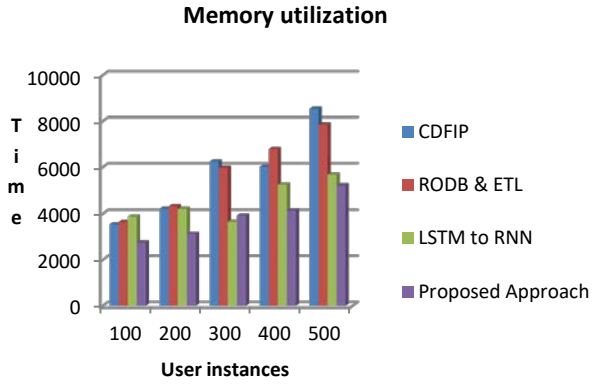

Fig. 4. Graph Showing Memory Utilization.

Fig. 4 shows a clear comparison of the reduced memory utilization by the system in comparison to the existing algorithms.

Hence the simulation results of the proposed algorithm outperform in terms of Performance, Encryption and Decryption throughput and memory utilization providing improved Privacy for patient sensitive data.

VI. CONCLUSION

This paper aimed to discuss the importance of security of patient data based on the access priorities of users, using a Hybrid ABE Approach. In due course several techniques related to mobile healthcare and e-healthcare grabbed concentration in research, but lacked profound architecture to preserve patient data. Our framework offers a innovative and

qualitative technique using SHA-1 and improved Feistel network in key generation ensuring authentication, and confidentiality during transmission entailing limited access to user communities considering access policy. The groups of users with limited access are subjected to Anonymity techniques. The result of our method renders improvised performance in several evaluation parameters considered. Lastly, we conclude that the roadmap presented endeavors a feasible solution for discussed privacy issues.

REFERENCES

[1] Supriya menon M & Rajarajeswari P, " A Hybrid Machine Learning approach for Drug Repositioning," IJCSNS International Journal of Computer Science and Network Security, VOL.20 No.12, December 2020, https://doi.org/10.22937/IJCSNS.2020.20.12.24.

[2] Saran, P., Rajesh, D., Pamnani, H., Kumar, S., Hemant Sai, T. G., & Shridevi, S, "A Survey on Health Care facilities by Cloud Computing, " International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE). doi:10.1109/ic-etite47903.2020.231.

[3] Chen, S.-W., Chiang, D. L., Liu, C.-H., Chen, T.-S., Lai, F., Wang, H., & Wei, W., " Confidentiality Protection of Digital Health Records in Cloud Computing.," Journal of Medical Systems, 40(5), 2016, doi:10.1007/s10916-016-0484-7.

[4] Zhiqiang, G., Lingsong, H., Hang, T., & Cong, L., " A cloud computing based mobile healthcare service system," 2015 IEEE 3rd International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), doi:10.1109/icsima.2015.7559009.

[5] Haufe, K., Dzombeta, S., & Brandis, K., " Proposal for a Security Management in Cloud Computing for Health Care,". The Scientific World Journal, 2014, pp. 1–7, doi:10.1155/2014/146970.

[6] Singh, I., Kumar, D., & Khatri, S. K, " Improving The Efficiency of E-Healthcare System Based on Cloud," 2019, Amity International Conference on Artificial Intelligence (AICAI), doi:10.1109/aicai.2019.8701387.

[7] Kumar, J., Goomer, R., & Singh, A. K., "Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model For Cloud Datacenters," Procedia Computer Science, 125,2018, pp. 676–682. doi:10.1016/j.procs.2017.12.087.

[8] Chiuchisan, I., Balan, D.-G., Geman, O., Chiuchisan, I., & Gordin, I., " A security approach for health care information systems," 2017, E-Health and Bioengineering Conference (EHB). doi:10.1109/ehb.2017.7995525.

[9] Masood, I., Wang, Y., Daud, A., Aljohani, N. R., & Dawood, H., "Towards Smart Healthcare: Patient Data Privacy and Security in Sensor-Cloud Infrastructure," Wireless Communications and Mobile Computing, 2018, 1–23. doi:10.1155/2018/2143897.

[10] Ronald Glasberg, Michael Hartmann, Michael Draheim, Gerrit Tamm, and Franz Hessel, " Risks and Crises for Healthcare Providers: The Impact of Cloud Computing," 2014, Academic Editors: R. Colomo-Palacios, M. Niedermayer, and V. Stantchev.

[11] Al-Issa, Y., Ottom, M. A., & Tamrawi, A., " eHealth Cloud Security Challenges: A Survey," Journal of Healthcare Engineering, 2019, 1–15. doi:10.1155/2019/7516035.

[12] Chenthara, S., Ahmed, K., Wang, H., & Whittaker, F., " Security and Privacy-preserving Challenges of e-Health Solutions in Cloud Computing," IEEE Access, 1–1. doi:10.1109/access.2019.2919982.

[13] Ramzi A. Haraty, Mirna Zbib and Mehedi Masud, "Data Damage Assessment and Recovery Algorithm from Malicious Attacks in HealthCare Data Sharing Systems," 2016, Secure cloud computing for mobile health services. Peer-to-Peer Networking and Applications, 9(5), 809–811. doi:10.1007/s12083-016-0451-6.

[14] Azeez, N. A., & der Vyver, C. V., " Security and privacy issues in e-health cloud-based system: A comprehensive content analysis," Egyptian Informatics Journal, 2018, doi:10.1016/j.eij.2018.12.001.

[15] Anitha, R., & Mukherjee, S., " Data Security in Cloud for Health Care Applications," Advances in Computer Science and Its Applications, 1201–1209, 2014, doi:10.1007/978-3-642-41674-3_167.

[16] Soceanu, A., Vasylenko, M., Egner, A., & Muntean, T., "Managing the Privacy and Security of eHealth Data," 2015, 20th International Conference on Control Systems and Computer Science. doi:10.1109/cscs.2015.76.

[17] Awan, I. A., Shiraz, M., Hashmi, M. U., Shaheen, Q., Akhtar, R., & Ditta, A., " Secure Framework Enhancing AES Algorithm in Cloud Computing," Security and Communication Networks, 2020, 1–16. doi:10.1155/2020/8863345.

[18] Opportunities and Challenges of Cloud Computing to Improve Health Care Services.

[19] Sun, P. J., " Privacy Protection and Data Security in Cloud Computing: A Survey, Challenges and Solutions," IEEE Access, 1–1. doi:10.1109/access.2019.2946185.

[20] Supriya menon M & Rajarajeswari P, "A contemporary way for enhanced modeling of context aware privacy system in PPDM,", Journal of Advanced Research in Dynamical and Control Systems.

[21] Uma Narayanan A , Varghese Paul B , Shelbi Joseph A , "A novel system architecture for secure authentication and data sharing in cloud enabled Big Data Environment," Engineering and Technology, Kochi, Kerala India.

[22] Ma, H., Zhang, R., & Yuan, W., Comments on "Control Cloud Data Access Privilege and Anonymity With Fully Anonymous Attribute-Based Encryption, " IEEE Transactions on Information Forensics and Security, 11(4), pp. 866–867,2016, doi:10.1109/tifs.2015.2509865.