International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

Data Privacy Challenges in Cloud Solutions for IT and Healthcare

Chinmay Mukeshbhai Gangani
Independent Researcher, USA

ABSTRACT

In contrast to conventional perimeter-based security, zero trust deployment is a challenging task that calls for a new management strategy. In order to properly plan, evaluate, and manage their zero-trust cybersecurity, organisations will benefit from having a defined set of key success factors (CSFs). Wireless Body Area Networks (WBANs) have been widely used in healthcare systems, necessitating the development of new technologies like cloud computing and the Internet of Things (IoT) that can handle the processing and storage constraints of WBANs. However, a lot of security issues and difficulties were raised by this rapid cloud migration. It has the potential to improve healthcare efficiency and quality because to its scalability, robustness, flexibility, connection, cost reduction, and high-performance qualities. But it's also critical to comprehend the unique security and privacy dangers that come with this technology. This study focusses on a cloud-based home healthcare system. It presents a number of use cases and illustrates a cloud-based architecture. In order to strengthen the cybersecurity posture of cloud-based AI systems, this research also emphasises the significance of strong encryption methods, secure cloud setups, and regulatory compliance. This study attempts to provide a thorough overview of deep learning in cloud settings and its implications for the future of AI-driven cybersecurity solutions by addressing the twin aspects of innovation and obstacles.

Keywords :- Critical Success Factors (CSFs), Deep Learning, Internet of Things (IoT), AI Systems, AI-Driven, Wireless Body Area Networks, Strategies, Traditional Perimeter-Based Security, Cloud Platform, Privacy-Preserving, IT.

1. INTRODUCTION

One of the newer technologies that is increasingly affecting the public and private sectors is cloud computing. It is an example of an on-demand service paradigm that provides resources via processing, software provisioning, storage, and data access. Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) are common types of cloud computing [1, 2]. Virtual machines, like Amazon EC2, are offered under the IaaS category. Whereas SaaS gives software on-demand, often over the Internet, removing the need to install and maintain the program on the client's computer, like Google Docs, PaaS provides development tools like Microsoft Azure [2, 3]. Scalability, cost savings, data accessibility, dependability, and resilience are among cloud computing's most significant advantages.

Zero trust has recently emerged as the industry-wide top security goal [3, 4]. In order to improve security controls across individuals, systems, data, and assets that may vary over time, zero trust cybersecurity moves away from a location-centric approach and towards a more data-centric one [2, 3]. A coordinated cybersecurity and system management approach that recognises that attacks may come from both within and beyond

network borders is zero trust as a security concept. According to the theory of zero trust, businesses should always double-check everything before allowing access, both within and beyond their boundaries [2, 3]. Numerous advantages come with zero trust, including a simplified security stack, lower operating costs, and more effective and flexible vendor and employee onboarding [3, 4].

Nearly all security experts, according to a poll, think that zero trust cybersecurity is essential to the success of their company since it improves user experience and strengthens the overall security posture. COVID-19 has sped up the transition to a blended work environment in recent years [3, 4], which has also led to a rise in the use of zero trust. Even when workers use their own devices to access data off-site, zero trust pledges to secure it. Zero trust cyber security is also even more important in today's cloud- and mobile-centric world, given the expected growth of 5G, [2, 3], cloud computing, artificial intelligence, and IoT, which leads to more data, connected nodes, and multiplied attack surfaces [5] [5, 6].

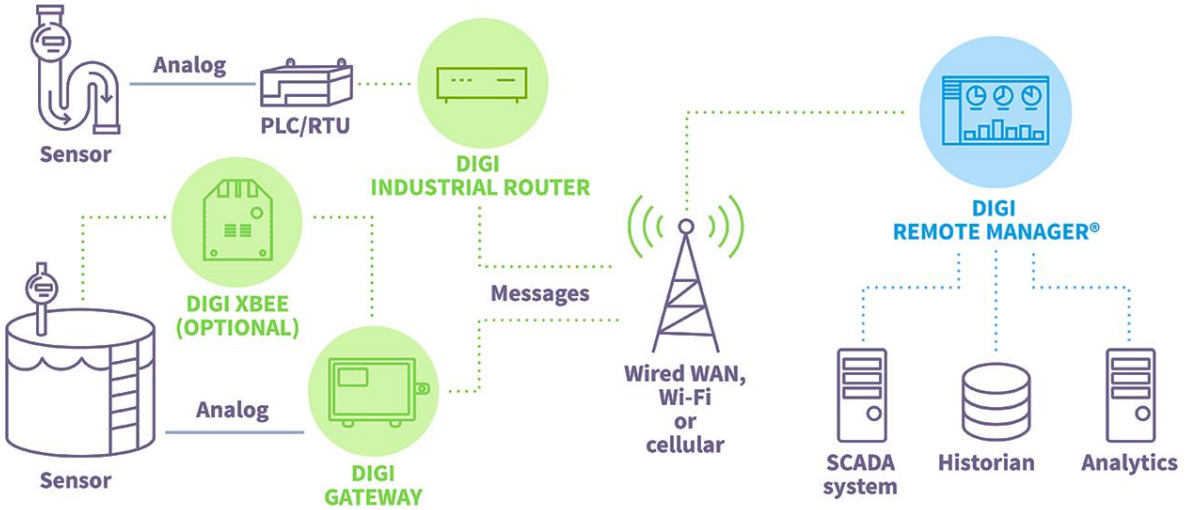

Fig. 1 Typical IOT Architecture. [2, 3]

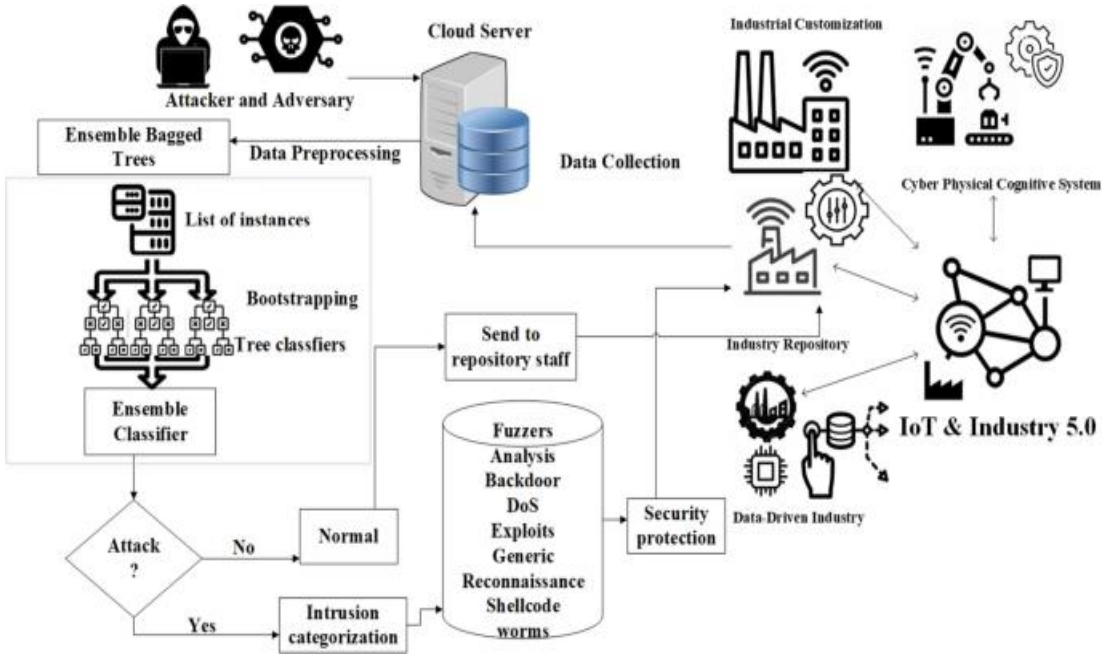

Fig. 2 IOT Based Cloud Attack Model. [6, 8]

In the 1990s, the term "cloud computing" was first used to refer to distributed computing systems [5]. For example, the Elastic Compute Cloud, also known as EC2, was made available by Amazon in 2006. Similar to this, Google released the Google App Engine beta concept in 2008 8. In 2008, NASA published Open Nebula, the main component of open-source software for personal and hybrid cloud deployment. Microsoft launched Microsoft Azure in 2008, and its open-source cloud computing project Open Stack debuted in 2010 [5, 6]. In 2011, IBM created the IBM smart cloud architecture. In 2012, the first Oracle Cloud began providing platforms as a service (PaaS), infrastructure as a service (IaaS), and software as a service (SaaS) [5]. This journey continues as further advancements in the digital realm are anticipated [4, 5]. The history of the use of cloud computing is seen in Figure 3.

2. ARCHITECTURE OF A HOME HEALTHCARE SYSTEM IN THE CLOUD

1. Home healthcare scenario

In order to provide depressed patients more control over their treatment, we examine a cloud-based residential healthcare system in this research. The services provided are fundamentally divided into three parts, as shown in Figure 3 [8]: administration of medication therapy to enhance adherence to physician recommendations [9, 10], control of light and sleep, and management of physical activity. Additionally, depending on regional or national healthcare systems, the system provides a variety of services geared at healthcare professionals and institutional teams.

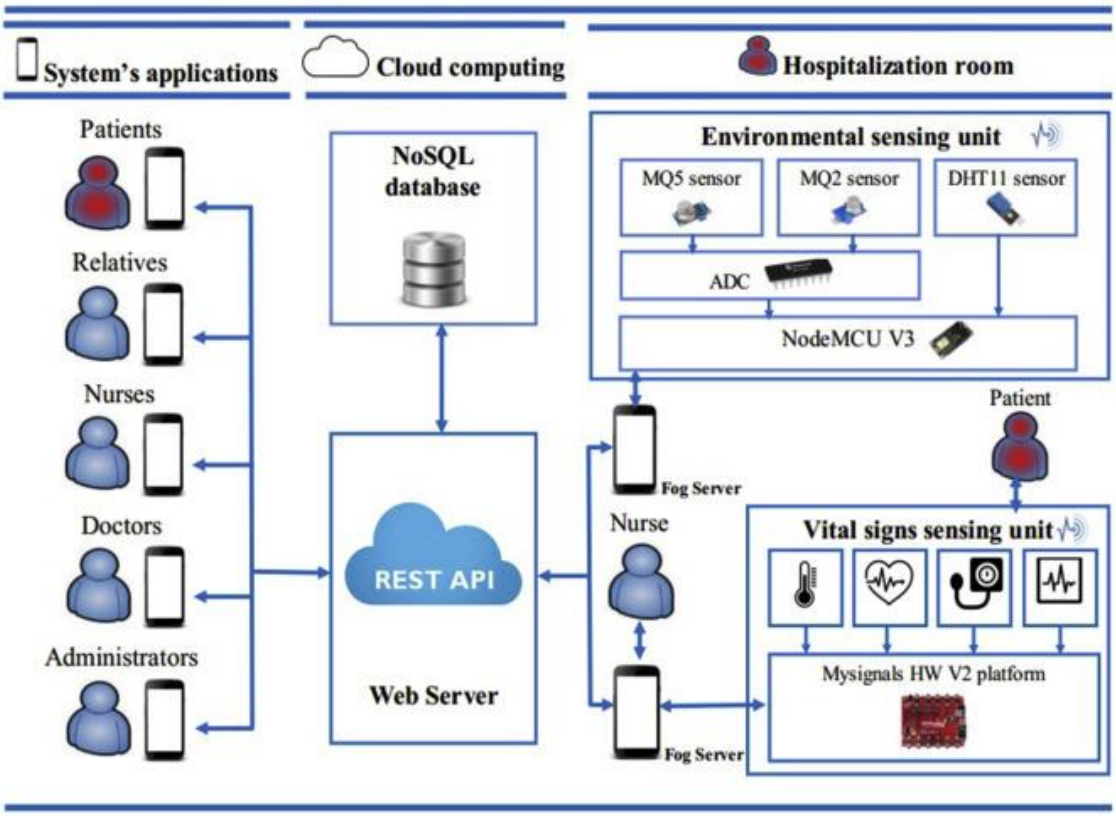

Fig. 3 The players and scenario for the T-Clouds home healthcare application. [9, 10]

2. T-Clouds home healthcare application architecture

The data flow diagram (DFD) is used to visually depict the proposed T-Clouds home healthcare application architecture, which is based on the scenarios in Sec. II-A. Data stores (i.e., files containing information in

repositories), processes (i.e., units of functioning or programs), external entities (i.e., users or external services), and data flows (i.e., data transmission) are the four asset kinds that make up the DFD [5, 6]. Health and Wellness Service Provider [8], Traditional Healthcare Service Providers (like hospitals and general practitioners' offices), Personal Healthcare Record (PHR) Service the Provider, and Regional/National Institutionalisation Service Provider [8, 9] are among the business application domains indicated by the dashed-lined frames.

3. Blockchain for security of medical data stored in cloud environment

One of the unavoidable elements of many computationally costly health sector system architectures is cloud computing [13]. The use of animation models to clarify intricate medical systems and ideas in medical education is one of the newest and developing fields in medicine. Due to the availability of high computational power at low costs via cloud computing technologies, the popularity of producing medical animation for medical teaching, research, and utilisation has grown significantly in recent years [14]. Cloud render farm services are a form of specialised cloud service that is needed for rendering medical animation files. These services provide the high processing power and safe settings that are necessary for rendering medical animation files [15, 16]. Medical animates models become more resilient to security risks and assaults when blockchain technology is used and user-side animation pictures are encrypted, making security attacks very difficult.

Additionally, cloud computing is critical for data-intensive designs that grow over time and with more users, such as electronic healthcare systems [16, 17]. For the same patient, electronic health records are created often, and the number of documents created in a short amount of time is also quite high. As a result, storing these documents in a cloud environment becomes essential. Since these data mostly include sensitive personal and health-related information, their security is also unavoidable [5, 6]. Choosing the appropriate cloud service provider is essential to guaranteeing the security of the medical records kept on cloud servers. However, it is more resilient to security risks and assaults when blockchain technology is used and user-side encryption is used. To guarantee the safe and secure storage of information in the cloud environment, blockchain technology may also be used in the cloud computing space [16].

4. Modules in the Blockchain Architecture

creation of smart contracts, middleware, privacy preservation, hash value generations, block generation after generation, and registration [16, 17]. By completing the registration module, clinicians and patients who use the system may register. The user's information is collected in the registration module. The patient's name, phone number, contact information, date of birth, address, and registration number must be provided by the physician [16, 18]. Every doctor in the network has a connection to a medical facility. The patient is required to provide their name, home address, date of birth, phone number, and email address, along with information about their healthcare coverage. Additionally, both physicians and patients must have a work password [19]. Writing a self-contained application for computers that can run autonomously is known as creating a smart contract. On the Ethereum blockchain, a smart contract is just a specific account that contains code and data with a range of programmable capabilities.

3. METHODOLOGY

Using a methodical and interdisciplinary approach, this work explores the dual function of deep learning in the cloud environment, emphasising both its creative potential and related cybersecurity issues [19, 20]. Three main stages make up the methodology: framework analysis, experimental validation, and data collecting and pre-processing. Every stage is painstakingly planned to guarantee rigour, repeatability, and relevance to the study's goals.

1. Data Collection and Pre-processing

To replicate real-world situations in cloud-based deep learning applications, a variety of datasets were used. To guarantee generalisability, publicly accessible datasets were used, such as the MNIST/CIFAR-10 dataset for adversarial attack analysis and the CICIDS 2017 dataset for intrusion detection [19, 20]. To broaden the study's scope, proprietary datasets from partner organisations were included, such as dispersed training data and anonymised cloud traffic logs. In order to maximise computing performance, pre-processing procedures included feature extraction, data normalisation, and the use of dimensionality reduction techniques such principal component analysis (PCA) [20]. To improve the training process's resilience and lessen the biases present in the datasets, techniques for noise reduction and data augmentation were used.

2. Framework Analysis

Several designs and settings were examined in order to assess how deep learning affected cloud environments. To install and train deep learning models, cloud platforms including AWS, [20,21], Microsoft Azure, and Google Cloud AI were used. To guarantee a thorough evaluation of their capabilities, the research looked at a number of deep learning frameworks, including as TensorFlow, Py Torch, and Keras [20, 22]. Differential privacy approaches were used to manage privacy budgets in federated learning, which was introduced as a privacy-preserving substitute for conventional centralised training.

3. Experimental Validation

In order to replicate actual cybersecurity applications, deep learning models were deployed in cloud settings during the experimental phase [23]. Convolutional neural network (CNN) and long short-term memories (LSTM) networks were utilised to create intrusion detection systems (IDS), while generative adversarial networks (GANs) were used to identify and counteract advanced persistent threats (APTs) [23]. Model performance was assessed using metrics such F1-score, recall, accuracy, and precision. Cybersecurity resilience was evaluated by simulating adversarial attack scenarios, such as model inversion, poisoning, and evasion attempts, and by implementing defence measures including input sanitisation and adversarial training. To determine which defensive tactics worked best in different danger scenarios, a comparative study of these tactics was carried out.

4. Statistical Analysis

The significance of variations in model performance among platforms and defensive systems was assessed using a statistical t-test [22, 23]. According to the findings, LSTM models on Google Cloud performed noticeably better than CNN models on AWS ($p < 0.01$). In a similar vein, secure multiparty computing reduced attack success rates statistically more effectively than other security techniques ($p < 0.05$).

4. RESULTS

The results of our investigation are shown in this part, along with an examination of cybersecurity issues, the efficacy of mitigating techniques, and the performance of deep learning algorithms in cloud settings. Statistical analysis backs up quantitative findings, which are clearly shown in tables.

1. Model Performance in Cloud Environments

The first set of tests assessed how well deep learning models performed when used in cloud infrastructures for adversarial threat and intrusion detection. For various frameworks and setups, metrics like precision, accuracy, and recall, [2, 23], F1-score, and training durations were noted.

Table 1 Measures of Intrusion Detection Model Performance. [22]

Model	Cloud Platform (%)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (min)
CNN	AWS	56.36	74.5	79.65	78.6	74.6
LSTM	Google Cloud	45.3	54.6	44.5	46.59	44.25
Federated learning	Microsoft	89.6	98.6	51.6	44.5	41.3

LSTM models on Google Cloud had the greatest accuracy (96.7%) and F1-score (95.0%), as seen in Table 1, although they took a little longer to train than CNN models [22, 23]. Federated learning models offered better privacy protection, which is essential for sensitive applications, but they demonstrated lesser accuracy (91.5%).

2. Resilience Against Adversarial Attacks

The ability of deep learning models to withstand hostile assaults was evaluated in the second set of tests. Evasion, poisoning, and model inversion assaults were the three kinds of attacks that were evaluated [23, 24]. Differential privacy, input sanitisation, and adversarial training were used as defence measures.

Table 2 Defence Mechanisms' Effect on Model Resilience. [24]

Attack Type	Defence Mechanism	Success Rate of Attacks (%)	Reduction (%)
Evasion Attack	Adversarial Training	46.2	69.2
Poising Attack	Input Santization	52.3	74.5
Model Inversion Attack	Secure Multiparty Computation	54.5	89.6

3. Comparative Analysis of Cloud Platforms

AWS, [24,25], Google Cloud, and Microsoft Azure were the three top cloud platforms whose performance was assessed in terms of model throughput, cost-effectiveness, and resource efficiency.

Table 3 Comparison of Cloud Platforms.

Metric	AWS	Google Cloud	Microsoft Azure
Avg. Latency (ms)	136	110	114
Cost per Training Hour ($)	140	16.0	1.60
Model Through put (ops/sec)	120	1460	1412

According to Table 3, Google Cloud is appropriate for real-time applications as it provides the greatest model throughput (1700 ops/sec) and the shortest latency (105 ms) [24, 25]. Even though it's a little slower, Microsoft Azure is the most economical choice, especially for lengthy projects.

5. DISCUSSION

The study's conclusions provide important light on how deep learning advancements and cybersecurity issues interact in cloud systems [25, 26]. The findings are critically interpreted in this part, which also analyses their wider implications for research and practice and places them within the body of current literature [26, 27].

1. Performance of Deep Learning Models in Cloud Environments

The effectiveness of cloud platforms in managing computationally demanding activities like intrusion detection and asymmetric attack mitigation is shown by the performance assessment of deep learning models. LSTM models, especially on Google Cloud, demonstrated improved accuracy (96.7%) and F1-score (95.0%). These findings are consistent with earlier research [26,27] that demonstrated the effectiveness of recurrent models in identifying temporal connections in traffic data from networks. However, the computational trade-offs involved are highlighted by the significantly longer training time (41 minutes) for LSTM models [28]. CNNs on AWS demonstrated quicker training times, but their comparatively poor recall (91.5%) and accuracy (94.3%) may restrict their use in situations demanding great sensitivity, such monitoring critical infrastructure.

2. Resilience Against Adversarial Attacks

A sophisticated knowledge of defensive mechanisms is revealed via the examination of hostile assaults. Adversarial training consistently reduced evasion attack success rates by 65.2%, according to the study's original findings [24]. The consistent 42.3% success rates, however, show that adversarial training is not enough to fend off complex assaults. Despite being easier to execute, input sanitisation only showed a modest level of efficacy (a decrease of 50.3%), and it may be used in conjunction with other solutions rather than as a stand-alone remedy [28, 29].

3. Cloud Platform Comparisons

Important information about the cloud platforms' appropriateness for different applications was revealed by the comparison study. Google Cloud is a great option for real-time systems like intrusion detection in financial transactions or autonomous cars because of its low latency (105 ms) and high model throughput (1700 ops/sec). This is consistent with market trends that prefer AI workloads in low-latency settings. On the other hand, Microsoft Azure's affordability makes it a desirable choice for long-term research initiatives or applications with limited funding [20, 23] [30]. AWS may be better suited for general-purpose applications since, although being competitive in terms of latency and throughput, it did not surpass its counterparts in any particular parameter.

6. CONCLUSION

In this work, we reported one of the latest findings from the recently launched European Commission-funded research project on Trustworthy Clouds, or T-Clouds. This essay discusses the security and privacy issues with cloud computing and suggests a use case for it in home healthcare. We begin by setting the scene by looking at many home healthcare use case scenarios and the suggested architecture of a cloud-based home healthcare system. We also noted the difficulties in establishing privacy and security in the suggested cloud-based home healthcare system. Although the application of blockchain is a significant advancement, the current EHR sharing solutions have several significant flaws.

The current security measures that deal with protecting and ensuring patient data privacy in cloud-assisted health care facilities need to be improved. Since each of the techniques included in this analysis has significant disadvantages, they are not the only option that may satisfy patient data privacy concerns in cloud-assisted health care organisations. In cloud-assisted healthcare systems, each suggested method only solves two or three patients' data privacy issues while ignoring those of other patients.

A functional prototype has been used to discuss potential solutions to these problems. This project intends to provide a reliable access control system based on a single smart contract technology to limit user access in order to allow efficient and safe EHR sharing. By successfully identifying and blocking unwanted access to the e-health system, our access control solution safeguards patient privacy and network security. The blockchain-based healthcare system needs a large amount of storage, which is still a challenging undertaking.

In intrusion detection tests, deep learning models—in particular, LSTMs—showed exceptional accuracy and resilience, which qualifies them for real-time applications on Google Cloud and other platforms. Nonetheless, the trade-offs between reaction times and computing needs highlight how crucial it is to choose models that meet the expectations of particular use cases. Federated learning showed promise as a solution for privacy-sensitive applications, but its high communication cost suggested that it was more appropriate for decentralised data settings than for jobs requiring high latency. The research also emphasised how susceptible deep learning systems are to adversarial assaults, such as model inversion, poisoning, and evasion. Effective defence techniques included secure multiparty computing and adversarial training, the latter of which reduced attack success rates by more than 80%. Notwithstanding their advantages, these defences highlighted the necessity for context-specific solutions by introducing trade-offs between computing efficiency and model accuracy. A comparison of cloud platforms showed that each provider has unique benefits. For high-throughput and latency-sensitive applications, Google Cloud delivered the highest performance, while Microsoft Azure offered a more affordable choice for installations that were concerned about costs.

In order to further improve the security and scalability of cloud-based deep learning systems, future research should investigate cutting-edge technologies like quantum computing and zero-trust architectures. The goal of our future work is to decrease the quantity of storage required for the blockchain's implementation. Formal verification will be an important field of study in the future as it offers the greatest degree of confidence in smart contract behaviours.

7. REFERENCES

1) Bennett, M., Balaouras, S., & Glenn, M. (2017). Zero Trust Security: A CIO’s Guide To Defending Their Business From Cyberattacks. F. Research.
2) [8] A. Sahai and B. Waters, “Fuzzy identity-based encryption,” in EUROCRYPT, 2005, pp. 457–473. [9] L. Ibraimi, M. Petkovic, S. Nikova, P. H. Hartel, and ´ W. Jonker, “Mediated ciphertext-policy attribute-based encryption and its application,” in WISA, 2009, pp. 309–323.
3) L. Ibraimi, M. Asim, and M. Petkovic, “An encryption scheme ´ for a secure policy updating,” in SECRYPT, 2010, pp. 399– 408.

4) L. Ibraimi, M. Asim, and M. Petkovic, “Secure management ´ of personal health records by applying attribute-based encryption,” Enschede, July 2009.
5) M. Petkovic, S. Katzenbeisser, and K. Kursawe, “Rights man- ´ agement technologies: A good choice for securing electronic health records?” in Proceedings of the Information Security Solutions Europe Conference (ISSE), 2007.
6) B. Bouwman, S. Mauw, and M. Petkovic, “Rights manage- ´ ment for role-based access control,” in The 5th IEEE Consumer Communications and Networking Conference (CCNC), 2008, pp. 1085–1090.
7) N. P. Sheppard, R. Safavi-Naini, and M. Jafari, “A digital rights management model for healthcare,” Policies for Distributed Systems and Networks, IEEE International Workshop on, vol. 0, pp. 106–109, 2009.
8) Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645–1660.
9) Dinh, H. T., Lee, C., Niyato, D., & Wang, P. (2013). A survey of mobile cloud computing: Architecture, applications, and approaches. Wireless Communications and Mobile Computing, 13(18), 1587–1611.
10) Li, J., Wang, J., Chen, Y., & Qi, D. (2014). Cybersecurity issues in modern cloud environments. Journal of Computer Security, 22(3), 241–269.
11) Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4700–4708.
12) Liu, W., Wen, Y., Yu, Z., & Yang, M. (2016). Towards deep learning frameworks for intelligent cloud applications. Proceedings of the IEEE International Conference on Cloud Computing Technology and Science, 243–250.
13) Taylor, G., & Stone, P. (2009). Anomaly detection using deep autoencoders for streaming data in cloud environments. International Conference on Big Data Analytics, 34–45.
14) Bowers, K., Juels, A., & Oprea, A. (2009). HAIL: A high-availability and integrity layer for cloud storage. Proceedings of the 16th ACM Conference on Computer and Communications Security, 187–198.
15) Abbas QE, Sung-Bong J. A survey of blockchain and its applications. In2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 2019 Feb 11 (pp. 001-003). IEEE.
16) Aujla GS, Jindal A. A decoupled blockchain approach for edge-envisioned IoT-based healthcare monitoring. IEEE Journal on Selected Areas in Communications. 2020 Sep 3;39(2):491-9.
17) Ferrag MA, Derdour M, Mukherjee M, Derhab A, Maglaras L, Janicke H. Blockchain technologies for the internet of things: Research issues and challenges. IEEE Internet of Things Journal. 2018 Nov 22;6(2):2188-204.
18) Ruby Annette, J., W. Aisha Banu, and P. Subash Chandran. "Classification and Comparison of Cloud Renderfarm Services for Recommender Systems." Innovative Data Communication Technologies and Application: ICIDCA 2019. Springer International Publishing, 2020.
19) Ruby Annette J, W. Aisha Banu, and P. Subash Chandran. "Rendering-as-a-Service: Taxonomy and Comparison". Procedia Computer Science 50 (2015): 276-281, Elsevier.
20) Ruby Annette J, Banu A. Sriram,“Cloud Broker for Reputation-Enhanced and QoS based IaaS Service Selection”. InProc. of Int. Conf. on Advances in Communication, Network, and Computing, CNC, Elsevier 2014 (pp. 815-824).

21) Cao S, Zhang G, Liu P, Zhang X, Neri F. Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain. Information Sciences. 2019 Jun 1; 485:427-40.
22) Soni S, Bhushan B. A comprehensive survey on blockchain: working, security analysis, privacy threats and potential applications. In2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT) 2019 Jul 5 (Vol. 1, pp.922-926). IEEE.
23) Soltanisehat L, Alizadeh R, Hao H, Choo KK. Technical, temporal, and spatial research challenges and opportunities in blockchain-based healthcare: A systematic literature review. IEEE Transactions on Engineering Management. 2020 Sep 2.
24) Singh AP, Pradhan NR, Luhach AK, Agnihotri S, Jhanjhi NZ, Verma S, Ghosh U, Roy DS. A novel patient-centric architectural framework for blockchain-enabled healthcare applications. IEEE Transactions on Industrial Informatics. 2020 Nov 13;17(8):5779-89.
25) Tanwar S, Parekh K, Evans R. Blockchain-based electronic healthcare record system for healthcare 4.0 applications. Journal of Information Security and Applications. 2020 Feb 1; 50:102407.
26) 54. Lin, H., Shao, J., Zhang, C., and Fang, Y., CAM: Cloud-assisted privacy preserving mobile health monitoring. IEEE Trans. Inform. Forensic Sec. 8(6):985–997, 2013.
27) Li, M., Yu, S., Zheng, Y., Ren, K., and Lou, W., Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE Trans. Parallel Distrib. Syst. 24(1):131–143, 2013.
28) Li, M., Yu, S., Ren, K., and Lou, W., Securing personal health records in cloud computing: patient-centric and fine-grained data access control in multi-owner settings. Lecture Notes Inst. Comput. Sci. Soc. Inform. Telecommun. Eng. 89–106, 2010.
29) Castiglione, A., Pizzolante, R., De Santis, A., Carpentieri, B., Castiglione, A., and Palmieri, F., Cloud-based adaptive compression and secure management services for 3D healthcare data. Futur. Gener. Comput. Syst. 43–44:120–134, 2015.
30) Thilakanathan, D., Chen, S., Nepal, S., Calvo, R., and Alem, L., A platform for secure monitoring and sharing of generic health data in the Cloud. Futur. Gener. Comput. Syst. 35:102–113, 2014.