IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

Image and Text Encrypted Data with Authorized Deduplication in Cloud

Sneha Solepatil[1], Snehal Divekar[2], Akanksha Nagare[3], Digvijay Gaikwad[4], Prof. M.S.Kale[5]
Department of Information Technology[1-5]
Sinhgad Academy of Engineering, Kondhwa, Pune, Maharastra, India

Abstract: *Cloud data storage is the most crucial service. Textual data and data pertaining to data bearers' privacy may occasionally be encrypted and kept in the cloud. Deduplication of encrypted text data continues to present new difficulties, and this has serious implications for the cloud's ability to store and handle large amounts of data. Cloud data should be protected so that unauthorised users can't access it. A data security technique called encryption is now available. The data is kept in the cloud in an encrypted manner to protect the users' privacy and security. Data access control and revocation cannot be flexible supported by these issues. We implement a method to deduplicate encrypted textual data saved in cloud-based services in this work.*

In order to prevent wasting the storage capacity supplied by cloud providers, good file storage and management is now crucial. The technique of data de-duplication, which only allows for the storage of one copy of a file, is commonly used to prevent file duplication in cloud storage systems. It contributes to a significant cost reduction for cloud service consumers by reducing the amount of storage space and bandwidth required. Data that must be stored today is encrypted for security purposes. Therefore, since data encryption with a key changes data, data encryption by data owners with their own keys prevents cloud service subscribers from doing data de-duplication. Cloud computing is a network-based computing system with a large storage area that allows authorized users to access the platform from any location at any time as long as there is strong network or internet connectivity. Cloud computing is mostly used to supply the device with shared hardware, software, and resources on demand. Instead of using a desktop, it functions like a remote server on the internet to store, manage, and process data. Therefore, compared to other local computers, the working time is quicker.

Keywords: Machine Learning, AES, MD5,Proxy re-encryption, Role authorized tree, Approved deduplication, Privacy leakage

I. INTRODUCTION

As social media grows in popularity and use, people are posting, sharing, and sending data in record numbers. The majority of software apps, social media sites, and businesses utilize cloud services to store their massive amounts of data. Files with the same content might be uploaded by the same or different users, causing the system to store the same files again and over, wasting the relatively costly storage space purchased from cloud service providers. Existing cloud storage company de-duplicate data to minimize wasting space, which benefits both themselves and their consumers. Deduplication may save backup storage requirements by up to 9095 percent and regular file system storage requirements by up to 68 percent. Encrypting the same files with different keys entered by users results in the generation of different cypher messages, even though the underlying plain text is the same. 1 As a result, classicalencryption fails in data de-duplication on encrypted files. However, encryption is expected to protect the security and secrecy of data. Previous de-duplication technologies, however, cannot guarantee the data's robustness. Furthermore, many de1duplication technologies require the data owner to all be brought online 1 in order to exchange a convergence key, therefore decryption cannot be performed just at time it is requested. Previous systems did not address storage server assaults and data retrieval in such attacks. In this research, 1 we propose a de-duplication method which is based on an erasure correction technique that splits the file into shards and distributes it over several cloud storage providers' servers. Even if only one of the servers is attacked by an intruder, the system can re-generate the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

original files using the remaining of repaired shards. Like a outcome, the system can guarantee the encrypted file's dependability and robustness.

II. RELATED WORK

In this chapter we are going to have an overview about how much time does it took to complete each task like-Preliminar0y Survey Introduction and Problem Statement, Literature Survey, Project Statement, Software Requirement and Specification, System Design, Partial Report Submission, Architecture Design.Implementation, Deployment, Testing, Paper Publish, Report Submission and etcetera. This chapter also gives focus on stakeholder list which gives information about project type, customer of the proposed system, user and project member who developed the system.S. Nagaprasad et al. [8], Ajay S. Ladkat et al. [9], S. L. Bangare et al. [10-15], K. Gulati et al. [16], P. S. Bangare et al. [17-18], Xu Wu et al. [19], V. Durga Prasad Jasti et al. [20], A. S. Zamani et al. [21], M. L. Bangare et al. [22] and S. Mall et al. [23] have proposed various research models which were referred here.

III. METHODOLOGY

AES

The United States government selected the Advanced Encryption Standard (AES) as a symmetric block cypher to safeguard sensitive data.

AES is used to encrypt sensitive data in hardware and software across the globe. It is crucial for government computer security, cyber security, and the safeguarding of electronic data.

When the National Institute of Standards and Technology (NIST) declared that a replacement for the Data Encryption Standard (DES), which was beginning to be vulnerable to brute-force attacks, was required, AES was put into development.

AES includes three block ciphers:

AES-128 uses a 128-bit key length to encrypt and decrypt a block of messages.
AES-192 uses a 192-bit key length to encrypt and decrypt a block of messages.
AES-256 uses a 256-bit key length to encrypt and decrypt a block of messages

What are the features of AES?

NIST specified the new AES algorithm must be a block cipher capable of handling 128-bit blocks, using keys sized at 128, 192 and 256 bits.

Other criteria for being chosen as the next AES algorithm included the following:

Security. Competing algorithms were to be judged on their ability to resist attack as compared to other submitted ciphers. Security strength was to be considered the most important factor in the competition.

MD5

With the help of the cryptographic hashing technique Message Digest Algorithm 5 (MD5), a string of any length can be converted into a value of 128 bits. Despite vulnerabilities being discovered, MD5 is still frequently
used. The most popular method for examining the consistency of files is MD5. SSH, SSL, and IP
Sec are just a few security applications and protocols that utilise it. Some applications enhance the MD5 algorithm by hashing the plaintext numerous times or by adding a salt value.

IV. SYSTEM ARCHITECTURE

To provide clients with options, we have examined their needs and suggested a number of CRM dashboard-based applications. Growing supply codes, enterprise-good judgement integration, and structure builds are all part of our plan. All of the mentioned features are converted into character sprints in accordance with the most recent approaches used in the agile improvement strategy. The execution of the task is broken down into three phases.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

Requirement accumulating Application improvement
Deployment and Hosting

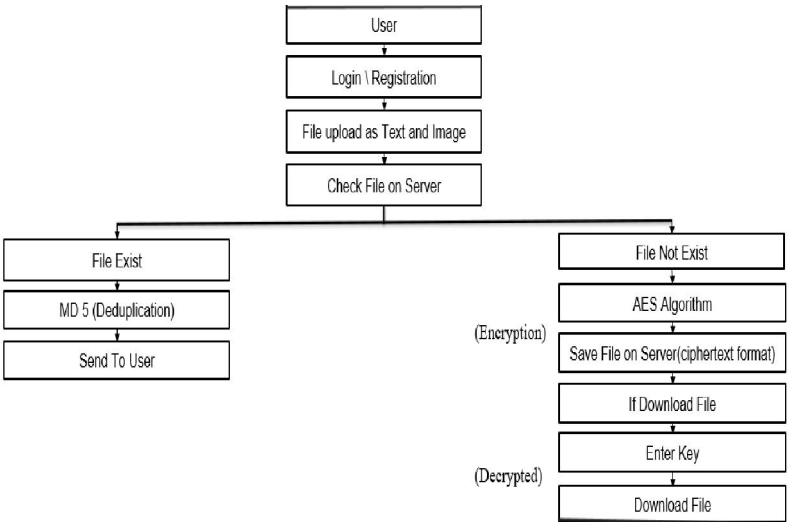


Fig.1. Architecture Diagram

A user sends a request to Management centre, and encrypts the file, then it results as the cipher text to CSP. Users who belong to completely different role groups owning the corresponding role keys, with the role keys the user access cloud server, the user will upload or transfer the files from Cloud Service Provider. And user can download the file from the cloud server. A Cloud Service Provider is mainly for data storage, management and verification. Cloud Service Provider stores and manages the uploaded files from authorized users. Management centre is the trusted third party that is for the authorized user and for the role key management.

V. ALGORITHM

AES - Encryption Decryption Algorithm

The AES encryption algorithm is a symmetric block cipher with a block size of 128 bits. It transforms these individuals block using 128, 192 and 256-bit keys. Once it has encrypted the blocks, it concatenates them to form the cipher text. Advanced Encryption Standard (AES) is a symmetric clock code developed in the United States. The government protects classified information. AES is implemented in software and hardware around the world to encrypt sensitive data. It is essential to government computer security, network security and electronic data protection.

MD 5 - Data Deduplication

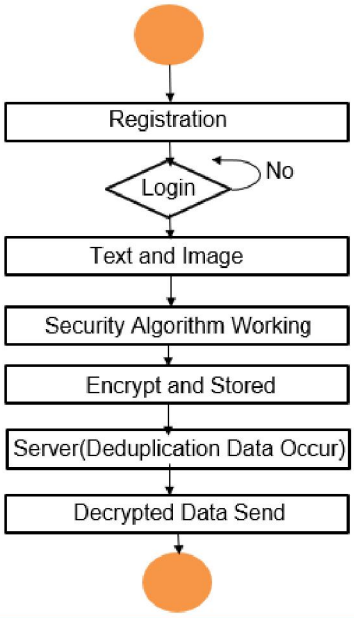
Methodology: -MD5: -

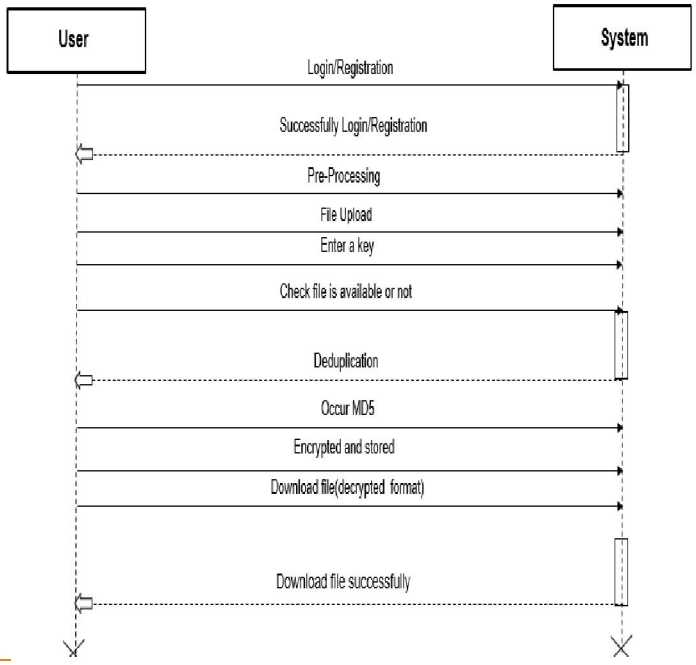
The MD5 message digest algorithm is a cryptographically broken but still commonly used hash function that produces a 128-bit hash value. The MD5 hash function was originally designed as a secure cryptographic hashing algorithm for verifying digital signatures. But MD5 has been decrypted for users other than as a non-cryptographic checksum verify a data integrity and detect unintentional data corruption.

VI. REQUIREMENT

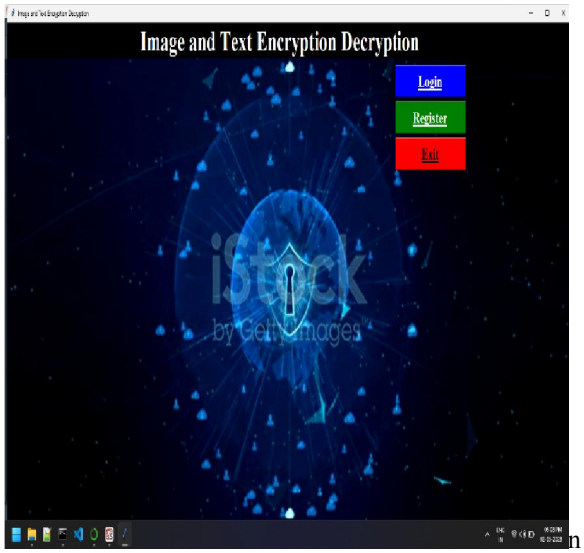
The performance of the functions and every module must be well. The overall performance of the software will enable the users to work efficiently. Performance of encryption of data should be fast. Performance of the providing virtual environment should be fast Safety Requirement. The application is designed in modules where errors can be detected and easily. This makes it easier to install and update new functionality if required.

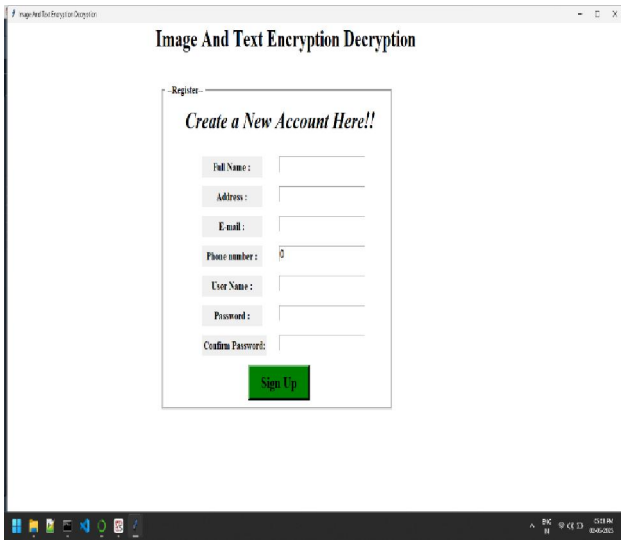
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023



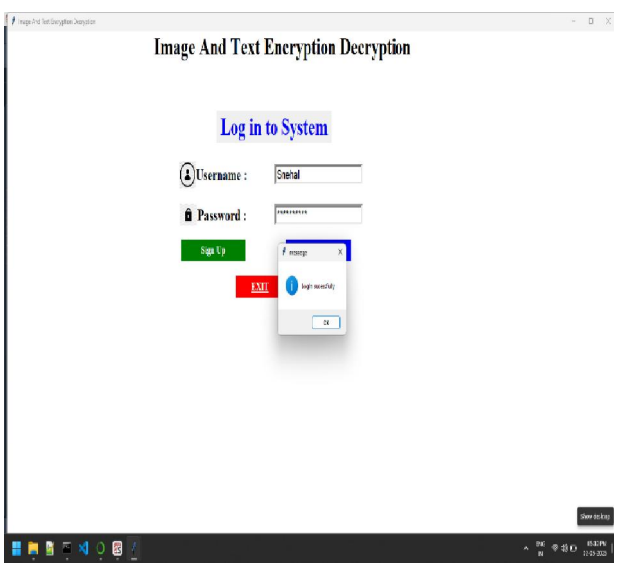


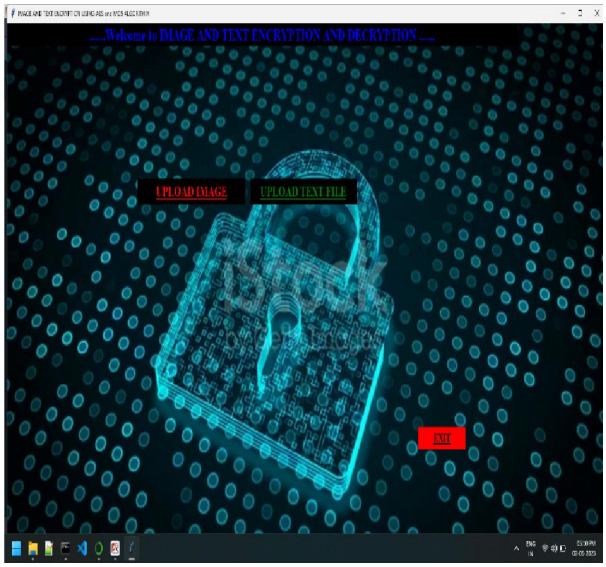
VII. RESULT AND DISCUSSION

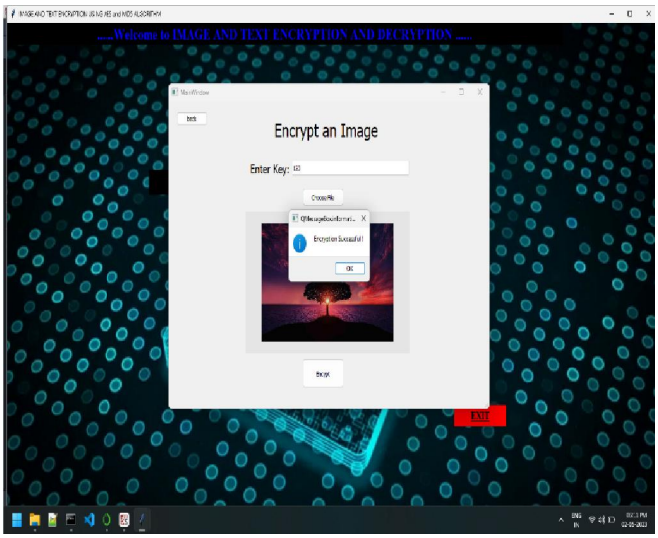




IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023







VIII. CONCLUSION

In this study, we discussed the use of encryption and decryption to prevent duplication. And we employ three algorithms for text uploading. We are utilizing the Structural Similarity AES Algorithm for cloud uploading, and the major goal of the similarity index is to examine the image quality, including luminance, contrast, and structure, before it calculates how similar two images are to one another. We use the encryption approach to store big amounts of data efficiently and prevent duplicating text and image.

ACKNOWLEDGEMENT

I hereby take this opportunity to record my sincere thanks and heartily gratitude to **Prof.M. S. Kale** for his useful guidance and making available to me his intimate knowledge and experience in making **"Image And Text Encrypted Data With Authorized Deduplication In Cloud"** as a preparation of report in respect thereof. I am also thankful to my HOD Dr. S.S.Kulkarni of my Information Technology department. I express my special thanks and heartily gratitude to my respective staff members for inspiring me throughout the completion of this system.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

REFERENCES

[1]. S. Halevi. D. Hornik. B. Pinkos. and A. Shulman-Peleg. "Proofs of ownership in remote storage systems,"" in Proceedings of the 18th ACM SIGSAC Conference on Computer and Communications Security. ACM, 20 I I, pp. 491-500

[2]. Gonzalez-Manzano and A. Orfila. "An efficient confidentialitypreselving proof of ownersh ip for deduplic,uion ," Journal of Network and Computer Applications. vol. 50, pp. 49-59, 2015.

[3]. J. Blasco, R. Di Pietro, A. Orfila, and A. Sorniotti. "A tunable proof of ownersh ip scheme for deduplication using bloom filters," in Comm\U1ications and Network Security (eNS). 2014 IEEE Conference on. IEEE.

[4]. W. K. Ng. Y. Wen, and H. Zhu, "Private data deduplication protocols in cloud storage," in Proceedin~ of the 27th Annual ACM Symposium on Applied Computinr; ACM, 20 12, pp. 441-446.

[5]. R. Oi Pietro and A. Sorniotti. "Boosting efficiency and security in proof of ownership for deduplicatioll." in Proceedings of the 7th ACM Symposium on Infonnation. Computer and Communications Security. ACM, 20 12, pp. 81-82.

[6]. M. Li. C. Qin, and P. P, C. Lee, "Cdstore; toward reliable, secure. and cost-efficient cloud storage via convergent dispersal," in Usenix Technical Conference, 20 15, pp. 45-53.

[7]. l. Xu, E.-C. Chang, and l. Zhou, "Weak leakage-resilient client-side deduplication of encrypted data.

[8]. S. Nagaprasad, D. L. Padmaja, Yaser Quereshi, S.L. Bangare, Manmohan Mishra, Mazumdar B. D., "Investigating the Impact of Machine Learning in Pharmaceutical Industry", Journal of Pharmaceutical Research International (Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759), Volume 33, Issue 46A, Pages 6-14, Publisher: JPRI https://www.journaljpri.com/index.php/JPRI/article/view/32834

[9]. Ajay S. Ladkat, Sunil L. Bangare, Vishal Jagota, Sumaya Sanober, Shehab Mohamed Beram, Kantilal Rane, Bhupesh Kumar Singh, "Deep Neural Network-Based Novel Mathematical Model for 3D Brain Tumor Segmentation", Computational Intelligence and Neuroscience, vol. 2022, Article ID 4271711, 8 pages, 2022. https://doi.org/10.1155/2022/4271711

[10]. S. L. Bangare, "Brain Tumor Detection Using Machine Learning Approach", Design Engineering ISSN: 0011-9342,Scopus Index- Q4, Ei Compendex, Volume 2021, Issue 7, Pages 7557-7566, Publisher Design Engineering.

[11]. S. L. Bangare, and P. S. Bangare. "Automated testing in development phase." International Journal of Engineering Science and Technology 4.2 (2012): 677-680.

[12]. S. L. Bangare, N. B. Dhawas, V. S. Taware, S. K. Dighe, & P. S. Bagmare, (2017). "Implementation of fabric fault detection system using image processing", International Journal of Research in Advent Technology, Vol.5, No.6, June 2017, E-ISSN: 2321-9637.

[13]. S. L. Bangare, N. B. Dhawas, V. S. Taware, S. K. Dighe, & P. S. Bagmare (2017). "Fabric fault detection using image processing method", International Journal of Advanced Research in Computer and Communication Engineering, 6(4), 405-409.

[14]. S. L. Bangare, S., H. Rajankar, P. Patil, K. Nakum, G. Paraskar, (2022). "Pneumonia detection and classification using CNN and VGG16". International Journal of Advanced Research in Science, Communication and Technology, 12, 771-779.

[15]. Sunil L. Bangare, Deepali Virmani, Girija Rani Karetla, Pankaj Chaudhary, Harveen Kaur, Syed Nisar Hussain Bukhari, Shahajan Miah, "Forecasting the Applied Deep Learning Tools in Enhancing Food Quality for Heart Related Diseases Effectively: A Study Using Structural Equation Model Analysis", Journal of Food Quality, vol. 2022, Article ID 6987569, 8 pages, 2022. https://doi.org/10.1155/2022/6987569

[16]. K. Gulati, M. Sharma, S. Eliyas, & Sunil L. Bangare (2021), "Use for graphical user tools in data analytics and machine learning application", Turkish Journal of Physiotherapy and Rehabilitation, 32(3), 2651-4451.

[17]. P. S. Bangare, Ashwini Pote, Sunil L. Bangare, Pooja Kurhekar, and Dhanraj Patil, "The online home security system: ways to protect home from intruders & thefts." International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN (2013): 2278-3075.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

[18]. P. S. Bangare, S. L. Bangare, R. U. Yawle and S. T. Patil, "Detection of human feature in abandoned object with modern security alert system using Android Application," 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), Pune, India, 2017, pp. 139-144, doi: 10.1109/ETIICT.2017.7977025.

[19]. Xu Wu, Dezhi Wei, Bharati P. Vasgi, Ahmed Kareem Oleiwi, Sunil L. Bangare, Evans Asenso, "Research on Network Security Situational Awareness Based on Crawler Algorithm", Security and Communication Networks, vol. 2022, Article ID 3639174, 9 pages, 2022. https://doi.org/10.1155/2022/3639174.

[20]. V. Durga Prasad Jasti, Enagandula Prasad, Manish Sawale, Shivlal Mewada, Manoj L. Bangare, Pushpa M. Bangare, Sunil L. Bangare, F. Sammy, "Image Processing and Machine Learning-Based Classification and Detection of Liver Tumor", BioMed Research International, vol. 2022, Article ID 3398156, 7 pages, 2022. https://doi.org/10.1155/2022/3398156

[21]. Zamani, A. S., Dr. Seema H. Rajput, Dr. Harjeet Kaur, Dr.Meenakshi, Dr. Sunil L. Bangare, & Samrat Ray. (2022). Towards Applicability of Information Communication Technologies in Automated Disease Detection. International Journal of Next-Generation Computing, 13(3). https://doi.org/10.47164/ijngc.v13i3.705.

[22]. M. L. Bangare, P. M. Bangare, R. S. Apare, & S. L. Bangare, (2021). "Fog computing-based security of IoT application", Design Engineering, 7, 7542-7549.

[23]. S. Mall, A. Srivastava, B. D. Mazumdar, M. Mishra, S. L. Bangare, & A. Deepak, (2022). "Implementation of machine learning techniques for disease diagnosis", Materials Today: Proceedings, 51, 2198-2201.https://www.sciencedirect.com/science/article/abs/pii/S2214785321072679#!